



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Analysis of O-island deletions in *Escherichia coli* O157:H7

Allen Forrest Flockhart



Thesis presented for the degree of Doctor of Philosophy

The University of Edinburgh

2012

Declaration

The research presented in this thesis is entirely my own work, except where otherwise stated. No part of this thesis has been submitted for a degree or professional qualification in any University.

Allen Forrest Flockhart

October 2012

Acknowledgments

First and foremost I would like to thank my supervisors, Professor David Gally and Dr Chris Low, for giving me the chance to do a PhD. This opportunity has been truly life changing for me and I will forever be indebted. I am also extremely grateful for all the advice and support they have given me throughout the duration of my studies. As my principal supervisor, I particularly thank Dave for his encouragement and for being a great mentor.

I would like to thank Tracy R, Tracy D, Fang, Jianing, Eliza, Dai, James, Makrina and Luke; for all of their friendships, advice and support over the years. I would also like to thank Alison and the LBEP girls, Jeanette, Bethan, Gill and Robyn; for their lab camaraderie and continued friendships; and I thank Sean for all his help and advice and for being the single best resource in the lab bar none!

I thank Jai, Tammie and Ronen for useful discussion and for kindly carrying out the gels shifts and IVT. I thank Maryia for sharing my enthusiasm for RgdR and for her incredible dedication with the fusions and RT-qPCR during the all too short summer project! I thank Yongxiang Zhang for helpful discussions; Dr

Darren Shaw for his help with statistical analyses; Professor Erik Denamur for kindly providing me with commensal *E. coli* strain ED1a; Professor Trinad Chakraborty for his generous supply of monoclonal antibodies; The VLA Weybridge for their expertise and help with the *in vivo* colonisation study; Dr John Leong and Dr Kenneth Campellone for their help constructing the OI deletion collection; and Dr Andrew Roe for his ideas and support during the early stages of this research. I would like to acknowledge the gift of purified H-NS and input on techniques from Dr Colin Corcoran, Prof. Charles Dorman and Prof. Jay Hinton; and I gratefully acknowledge The Royal (Dick) School of Veterinary Studies for funding my research.

I would like to give special thanks to Neil for his support throughout. My paper and thesis could not have happened without you. I also give special thanks to my 'bairns', Boots and Olive, for their companionship, love and affection; and for basically keeping me going!

Finally, I would like to proudly acknowledge my mum and granny whom unfortunately could not be here to see me finish my PhD, but whom have always loved and supported me.

Contents

Title page	i
Declaration	ii
Acknowledgments	iii
Contents	v
List of abbreviations	xi
List of figures	xvii
List of tables	xix
Abstract	xx
Chapter 1. Introduction	1
1.1 Introduction to <i>E. coli</i>	2
1.2 Pathotypes of <i>E. coli</i>	2
1.2.1 Enterotoxigenic <i>E. coli</i> (ETEC)	3
1.2.2 Enteroinvasive <i>E. coli</i> (EIEC)	6
1.2.3 Enteropathogenic <i>E. coli</i> (EPEC)	7
1.2.4 Enteroaggregative <i>E. coli</i> (EAEC)	9
1.2.5 Enterohaemorrhagic <i>E. coli</i> (EHEC)	11
1.2.5.1 Nomenclature	11
1.2.5.2 Clinical importance	12
1.2.5.3 Epidemiology	14
1.2.5.4 Reservoirs	14
1.2.5.4.1 Cattle	14
1.2.5.4.2 Non-ruminants	16

1.2.5.5	Transmission	16
1.2.5.5.1	Food-borne	16
1.2.5.5.2	Water-borne and environmental	18
1.2.5.5.3	Person to person	19
1.2.5.5.4	Incidence	20
1.3	Pathogenesis and virulence of EHEC O157:H7	22
1.3.1	Shiga-toxin	22
1.3.1.1	Mode of action	22
1.3.1.2	Role in intestinal disease	24
1.3.1.3	Role in HUS	25
1.3.2	Attaching and effacing (A/E) lesion formation	26
1.3.2.1	Role of locus enterocyte effacement (LEE) in A/E lesion formation	27
1.3.2.2	LEE and type III secretion (T3S)	28
1.3.2.3	Regulation of LEE and T3S	32
1.3.2.4	Intrinsic regulation of LEE	33
1.3.2.5	Horizontally acquired regulators of LEE	37
1.3.2.6	<i>E. coli</i> regulators of LEE	39
1.3.2.7	Environmental regulators of LEE	41
1.3.2.8	Post-transcriptional and post-translation control of LEE	45
1.3.2.9	Non-LEE encoded effector proteins	48
1.3.3	Plasmid encoded factors	51
1.3.3.1	Enterohaemolysin	52
1.3.3.2	EspP	53
1.3.3.3	EHEC type II secretion system (T2S)	54

1.3.3.4	KatP	54
1.3.3.5	StcE	55
1.3.3.6	ToxB	56
1.3.4	Fimbrial adhesins	57
1.3.4.1	Type I fimbriae	57
1.3.4.2	Long polar fimbriae	58
1.3.4.3	Curli	59
1.3.4.4	Type IV pili	61
1.3.5	Non-fimbrial adhesins	62
1.3.6	Flagella	63
1.4	Evolution of EHEC	66
1.5	Comparative genomics: EHEC O157:H7 versus K12	67
1.6	Research aims and rationale	69
 Chapter 2. Materials and methods		 72
2.1	Bacterial strains and plasmids	73
2.2	Bacterial culture conditions and media	73
2.3	DNA analysis and genetic manipulation methods	74
2.3.1	Genomic DNA extraction from <i>E. coli</i> strains	74
2.3.2	Preparation of crude DNA extracts from <i>E. coli</i> strains	75
2.3.3	Plasmid DNA extraction from <i>E. coli</i> strains	75
2.3.4	DNA amplification by PCR	76
2.3.5	DNA electrophoresis	76
2.3.6	DNA fragment purification	77
2.3.6.1	Column and gel purification	77

2.3.7	Transformation of competent <i>E. coli</i>	78
2.3.8	DNA digestion	78
2.3.9	Klenow-fill in	78
2.3.10	DNA ligation	79
2.3.11	Preparation and transformation of chemically-competent <i>E. coli</i> cells	79
2.3.12	Preparation and transformation of electro-competent <i>E. coli</i> cells	81
2.3.13	TOPO TA cloning	82
2.3.14	Construction of OI deletions in EHEC strain TUV93-0	83
2.3.15	Construction of defined mutations in EHEC and UPEC	83
2.3.16	Construction of plasmids for complementation	84
2.3.16.1	Construction of engineered ECs1581 and ECED1_1787 protein variants	84
2.3.17	Measurement of <i>LEE1</i> promoter activity	84
2.3.18	Reverse transcription-quantitative PCR (RT-qPCR)	85
2.4	Bioinformatic analyses	86
2.5	Sequencing	87
2.6	Growth rate in M9 minimal media	87
2.7	Acid survival	87
2.8	Congo-red binding	88
2.9	Motility	89
2.10	<i>In vivo</i> colonisation model	89
2.11	Methods for protein analysis	90
2.11.1	Preparation and analysis of T3S culture supernatant proteins	90

2.11.2	Sodium-dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)	92
2.11.3	Colloidal blue staining of SDS-PAGE gels	92
2.11.4	Protein detection	93
2.11.5	Native protein purification	94
2.12	Electrophoretic mobility shift assay (EMSA)	96
2.13	<i>In vitro</i> transcription (IVT)	96
2.14	Data analyses	97
Chapter 3.	Phenotypic screening of OI deletions	115
3.1	Growth analysis in M9 minimal media	116
3.2	Acid survival	118
3.3	Motility	120
3.4	Analysis of T3S culture supernatant proteins	122
3.5	Discussion	124
Chapter 4.	Genotypic and phenotypic characterisation of OI-51	131
4.1	<i>In vivo</i> analysis of OI-51	132
4.2	OI-51 sequence analysis	134
4.3	OI-51 cloning and complementation	136
4.4	Analysis of cloned OI-51 fragment z1835-z1843	137
4.5	Identification of <i>ecs1581</i> as a positive regulator of T3S	143
4.6	Discussion	145

Chapter 5. Functional and mechanistic characterisation of ECs1581	149
5.1 Distribution of ECs1581 in <i>E. coli</i> strains	150
5.2 Analysis of ECs1581 orthologous protein C1493	154
5.3 Analysis of ECs1581 activity in a <i>grlA</i> mutant	155
5.4 Analysis of ECs1581 orthologous protein ECED1_1787	157
5.5 Analysis of natural and engineered variants of ECED1_1787 and ECs1581	160
5.5.1 Analysis of EHEC lineage specific variants of ECs1581	161
5.5.2 Analysis of Congo-red binding activity	163
5.6 Analysis of ECs1581 DNA binding activity	167
5.7 Analysis of ECs1581 transcription activity <i>in vitro</i> (IVT)	169
5.8 Analysis of ECs1581 mechanism of action	171
5.9 Discussion	177
 Chapter 6. General discussion	 189
6.1 General discussion	190
6.1.1 <i>E. coli</i> O157:H7: current understandings in relation to this research	190
6.1.1.1 Summary of research findings	193
6.1.1.2 OI-51 encoded RgdR proteins are present in diverse phylogenetic groups of <i>E. coli</i>	196
6.1.1.3 Future work to address the functionality of OI-51-like prophage in <i>E. coli</i>	201
6.1.1.4 Future perspective	209

List of abbreviations

A/E	Attaching and effacing
AGR	Adjacent genomic regions
Amp ^R	Ampicillin resistance
APS	Ammonium persulfate
ATP	Adenosine triphosphate
BAX	BCL2-associated X protein
Bfp	Bundle forming pili
BLAST	Basic local alignment sequence tool
bp	Base pairs
BSA	Bovine serum albumin
cAMP	cyclic adenosine-3',5'-monophosphate
Cam ^R	Chloramphenicol resistance
CFA	Colonisation factor antigen
cfu	Colony forming units
CP	Cryptic prophage
CR	Congo red
C-terminus	Carboxy-terminus
Dam	DNA adenine methylase
dH ₂ O	Distilled water
DMEM	Dulbecco's minimal essential media
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotise triphosphate
DTT	Dithiothreitol

<i>E. coli</i>	<i>Escherichia coli</i>
EAEC	Enterotoxigenic <i>E. coli</i>
EDTA	Ethylenediamine tetraacetic acid
EHEC	Enterohaemorrhagic <i>E. coli</i>
Ehx	Enterohaemolysin
EIEC	Enteroinvasive <i>E. coli</i>
EMSA	Electrophoretic mobility shift assay
EPEC	Enteropathogenic <i>E. coli</i>
Esp	<i>E. coli</i> secreted protein
ETEC	Enterotoxigenic <i>E. coli</i>
ExPEC	Extra-intestinal pathogenic <i>E. coli</i>
F primer	Forward primer
Fis	Factor for inversion sequence
g	gram
gfp	Green fluorescent protein
GI	Gastrointestinal tract
HC	Haemorrhagic colitis
HCL	Hydrochloric acid
HCP	Haemorrhagic colitis pilus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HGT	Horizontal gene transfer
H-NS	Histone-like nucleoid structuring protein
HRP	Horseradish peroxidase
HUS	Haemolytic uraemic syndrome
IHF	Integration host factor

IL-	Interleukin
IPTG	Isopropyl β -D-1-thiogalactopyranoside
JNK	c-Jun N-terminal kinase
Kan ^R	Kanamycin resistance
Kb	Kilobase
<i>K_d</i>	Dissociation constant
kDa	Kilodalton
LB	Luria Broth
LEE	Locus of enterocyte effacement
Ler	LEE-encoded regulator
LPF	Long polar fimbriae
M	Molar
Map	Mitochondria-associated protein
Mb	Megabase
M-cell	Membranous cell
MDa	Megadalton
μ g	Microgram (10^{-3} gram)
μ l	Microlitre (10^{-3} litre)
μ M	Micromolar (10^{-3} molar)
MEM	Minimal essential media
mg	Milligram (10^{-3} gram)
ml	Millilitre (10^{-3} litre)
MLEE	Multilocus enzyme electrophoresis
MLST	Multilocus sequence typing
mM	Millimolar (10^{-3} molar)

Nal ^R	Nalidixic acid resistance
Nck	Non-catalytic region of tyrosine kinase adaptor protein
NEB	New England Biolabs
Nle	Non-LEE encoded proteins
NMEC	Newborn meningitis <i>E. coli</i>
N-terminus	N-terminus
N-WASP	Neuronal Wiskott-Aldrich syndrome protein
OD ₆₀₀	Optical density at 600 nm
OI(s)	O-island(s)
OM	Outer membrane
ORF	Open reading frame
P	Promoter
PAI(s)	Pathogenicity island(s)
PBS	Phosphate buffered saline
Pch	Plasmid encoded regulator C homologue
PCR	Polymerase chain reaction
Per	Plasmid encoded regulator
PerC	Plasmid encoded regulatory protein C
pI	Isoelectric point
pO157	Plasmid O157
qRT-PCR	Quantitative reverse transcription PCR
QS	Quorum sensing
R primer	Reverse primer
RNA	Ribonucleic acid
rpm	Revolutions per minute

rRNA	Ribosomal RNA
RTX	Repeat in toxin
s.e.	Standard error of the mean
SDS	Sodium-dodecyl sulphate
SDS-PAGE	Sodium-dodecyl sulphate-polyacrylamide gel electrophoresis
SOC	Salt-optimised and carbon
Sp	Sakai prophage
STEC	Shiga toxin producing <i>E. coli</i>
Stx	Shiga toxin
T3S	Type III secretion
T3SS	Type III secretion system
Ta	Annealing temperature
<i>TaqI</i>	Restriction endonuclease I encoded by <i>Thermus aquaticus</i>
TBE	Tris-borate buffer
TCA	Trichloroacetic acid
TE	Tris EDTA buffer
TEMED	N,N,N',N'-Tetramethylethylene diamine
TFB	Transformation buffer
Tir	Translocated intimin receptor
TLR(s)	Toll-like receptor(s)
Tris	Trishydroxymethylaminomethane
U	Enzyme unit
UPEC	Uropathogenic <i>E. coli</i>
UTI	Urinary tract infection
V	Volt

v/v	Volume per volume
VLA	Veterinary Laboratories Agency
VTEC	Vero-toxin producing <i>E. coli</i>
wHtH	winged helix-turn-helix
wt	Wild type
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Yop	Yersinia outer protein
α	alpha
β	Beta
γ	Gamma

List of figures

Figure 1.1	Locus of enterocyte effacement (LEE) encoding a type III secretion system (T3SS).	30
Figure 1.2	Schematic summary showing the complex regulation of the LEE PAI in EHEC O157:H7	34
Figure 3.1	Growth phenotypes for EHEC O157:H7 strain TUV93-0 and a selection of sixteen isogenic OI deletion cultured in M9 minimal media.	117
Figure 3.2	Acid survival phenotypes for EHEC O157:H7 strain TUV93-0 and a selection of twelve isogenic OI deletions growth in LB broth pH 2.0.	119
Figure 3.3	Analysis of motility in EHEC O157:H7 strain TUV93-0 and a selection thirty nine isogenic OI deletions.	121
Figure 3.4	Screening of EHEC O157:H7 OI deletions for altered T3S levels.	123
Figure 4.1	Contribution of OI-51 to EHEC colonisation in sheep.	133
Figure 4.2	Schematic representation of the genomic organisation of OI-51 and similar prophages in a subset of <i>E. coli</i> and <i>Shigella</i> strains.	135
Figure 4.3	Analysis of T3S levels in an OI-51 deletion <i>in trans</i> expressing cloned OI-51 fragments.	139
Figure 4.4	Construction of pWZ1835-43 cut-back derivatives and the analysis of their effects on T3S levels.	140

Figure 4.5	Identification of <i>ecs1581</i> as a positive regulator of T3S.	144
Figure 5.1	Sequence comparison of ECs1581 with LEE-encoded regulator (Ler).	151
Figure 5.2	Sequence comparison of ECs1581 variants in <i>E. coli</i>	153
Figure 5.3	ECs1581 and orthologous protein C1493 can activate T3S and down-regulate motility in EHEC and UPEC	156
Figure 5.4	ECs1581 can stimulate T3S independently of GrlA	158
Figure 5.5	Effects of natural and engineered variants of ECs1581 on T3S and motility	159
Figure 5.6	EHEC O157:H7 lineage specific variants of ECs1581 all have the capacity to stimulate T3S and repress motility	165
Figure 5.7	ECs1581 enhances binding to Congo-red	166
Figure 5.8	ECs1581 interacts with the <i>LEE1</i> promoter in EHEC	168
Figure 5.9	ECs1581 represses <i>LEE1</i> transcription in an <i>in vitro</i> transcription set up.	170
Figure 5.10	ECs1581 stimulates T3S via <i>LEE1</i> .	173
Figure 5.11	Schematic model/interpretation of how ECs1581/RgdR acts to positively regulate <i>LEE1</i> and T3S in EHEC	175

List of tables

Table 2.1	Bacterial strains	98
Table 2.2	Bacterial plasmids	103
Table 2.3	Primers	108
Table 5.1	List of ECs1581 orthologous and paralogous proteins present in <i>E. coli</i> and <i>Shigella spp</i> strains	185

Abstract

Escherichia coli (*E. coli*) are a diverse species of bacteria that reside, often harmoniously and beneficially, in the gastrointestinal tracts of humans and other mammals. However, some strains are associated with serious intestinal and extra-intestinal disease and are considered pathogens. The main differences between strains of these different *E. coli* pathotypes can be explained by the acquisition of genetic information introduced by mobile genetic elements, in particular bacteriophage. In enterohaemorrhagic *E. coli* (EHEC) O157:H7 strain EDL933, a pathotype of *E. coli* containing prophage-encoded Shiga toxins associated with severe gastrointestinal and systemic disease in humans, these horizontally acquired elements have been termed O-islands (OIs) and include both fully functional and cryptic prophages. The overall aim of this research was to try and determine what these OIs are actually doing for the bacteria. Systems pertinent in the life cycle and virulence of this pathogen were therefore investigated by phenotypically screening a large library of OI deletions in EHEC strain TUV93-0, a Shiga toxin-negative derivative strain of EDL933, and then comparing these with the parent strain. These analyses highlighted a subset of OIs with the potential to regulate motility and type III secretion (T3S), the latter being an essential colonisation factor for EHEC that is encoded by the locus of

enterocyte effacement (LEE). Deletion of OI-51, a 14.93 Kb cryptic prophage designated as CP-933C, significantly reduced persistence of faecal shedding in sheep and levels of T3S expression *in vitro*. Cloning and complementation together with targeted allelic replacements in OI-51 identified a novel positive regulator of the LEE, encoded by *ecs1581* in the sequenced *E. coli* O157:H7 strain Sakai that is present but not annotated in the EDL933 sequence. Functionally important residues of ECs1581 were identified by site-directed mutagenesis based on phenotypic variants present in strains from different *E. coli* pathotypes, including strains not harbouring a LEE-encoded T3S system. This regulator was subsequently termed RgdR based on a motif demonstrated to be important for stimulation of gene expression from *LEE1*. Purified RgdR protein was able to form multiple complexes on a PCR generated *LEE1* promoter fragment, and activation of this operon appeared to require this DNA binding capacity as a non-T3S inducing variant was unable to bind this same *LEE1* promoter fragment. RgdR did not directly activate *LEE1* transcription *in vitro*, nor did it activate transcription by relieving H-NS repression as proposed for the global regulator Ler (LEE-encoded regulator). However, RgdR activation did require a wild type *LEE1* promoter and the Ler auto-induction cycle to induce *LEE2-5* expression and T3S. RgdR was able to increase binding to Congo red and was

capable of repressing bacterial motility. Further analyses demonstrated that RgdR did not regulate T3S and cell motility via GrlA (global regulator of LEE activator) and QseC (quorum sensing *E. coli* regulator C), two established regulators in *E. coli* that control LEE gene expression and motility in conjunction with their partners, GrlR (global regulator of LEE repressor) and QseB (quorum sensing *E. coli* regulator B) respectively. RgdR is therefore identified as a novel regulator able to co-ordinate T3S and motility expression. This research has identified OI-51 as being important for EHEC O157:H7 colonisation in sheep and has identified a completely new family of small bacterial regulators that control surface factor expression in *E. coli*.

Chapter 1

Introduction

1.1 Introduction to *Escherichia coli*

Escherichia coli (*E. coli*) is a highly diverse and adaptable species of Gram-negative bacilli in the *Enterobacteriaceae* family. As the predominant facultative anaerobe of the mammalian colonic flora, *E. coli* typically colonises the gastrointestinal (GI) tract within hours of life and with the commensal flora plays a crucial role in developing and maintaining intestinal physiology (Nataro and Kaper, 1998). However, a subset of *E. coli* strains are associated with distinct syndromes of diarrhoeal disease, and different pathotype of this bacterium can also cause urinary tract infections, septicaemia and meningitis (Levine, 1987; Nataro and Kaper, 1998). The main differences between strains of these different *E. coli* pathotypes can largely be explained by the acquisition of genetic information introduced by mobile genetic elements, in particular, bacteriophage.

1.2 Pathotypes of *E. coli*

As mentioned, several distinct pathogenic categories of diarrheagenic *E. coli* are now recognised. Each pathotype is defined by a characteristic set of virulence-associated factors that in turn determine the clinical, pathological and

epidemiological features of the disease they cause. On the basis of these virulence associated determinants, diarrheagenic *E. coli* have been characterised into five principal pathotypes (Robins-Browne and Hartland, 2002). Serologically, each of these *E. coli* pathotypes can be distinguished by their distinctive O serogroups and O:H serotypes (Robins-Browne and Hartland, 2002). The current scheme for this classification of *E. coli* is a modified version of that originally proposed by the Danish bacteriologist Kauffman, in the 1940's. In accordance with this scheme, *E. coli* are serotyped on the basis of their surface antigens, namely O (somatic), H (flagellar) and K (capsular) (Kauffman, 1947; Nataro and Kaper, 1998). Although not without its limitations, this typing scheme has played an important role in the studies of the epidemiology and elucidating the pathogenesis of *E. coli* infection, and continues to do so though the scheme is now complemented by more modern typing techniques, such as multilocus sequence typing (MLST) that offer greater discrimination (Robins-Browne and Hartland, 2002).

1.2.1 Enterotoxigenic *E. coli* (ETEC)

ETEC are one of the major etiological agents of infantile and traveller's diarrhoea in less developed countries, and members of this group of

diarrheagenic *E. coli* are known to secrete at least one of two varieties of enterotoxin: (a) heat-stable enterotoxin (ST) and (b) heat-labile enterotoxin (LT) (Robins-Browne and Hartland, 2002). It is the action of these enterotoxins, described below, that causes the diarrhoeal disease seen following infection by ETEC strains. ETEC infection is acquired mainly by the ingestion of contaminated food and water; with the bacteria colonising the proximal small intestine (Levine, 1987). Clinical symptoms include loose stools, nausea, vomiting and abdominal cramps (Clarke, 2001).

(a) Heat-stable enterotoxin (ST)

ST is a low molecular weight peptide comprising 19 amino acids and shares significant homology with the intestinal paracrine hormone guanylin (Robins-Browne and Hartland, 2002). The heat stability of this monomeric enterotoxin is conferred by its multiple cysteine residues and their disulfide bonds (Nataro and Kaper, 1998). Two unrelated classes of ST exist, namely STa (or ST-I) and STb, both of which are plasmid mediated (Clarke, 2001). Genes for ST have also been found on transposons and in addition to ETEC, STa toxins are also produced by several other Gram negative bacteria, including *Yersinia*

enterocolitica (*Y. enterocolitica*) and *Vibrio cholerae* (*V. cholerae*) non-O1 (Nataro and Kaper, 1998). The major receptor for STa is a membrane-spanning enzyme called guanylate cyclase C (GC-C), located in the apical membrane of intestinal epithelial cells (Nataro and Kaper, 1998). Binding of STa to GC-C results in increased intracellular enzymatic activity and increased cyclic guanylate monophosphate (cGMP) levels (Nataro and Kaper, 1998), which in turn leads to the stimulation of chloride secretion and/or inhibition of sodium adsorption and subsequent intestinal fluid secretion (Nataro and Kaper, 1998). Although STb is primarily associated with ETEC infections in pigs, some human ETEC isolates express this toxin (Nataro and Kaper, 1998). In contrast to STa however, the receptor for STb is still unknown. This is largely due to the tissue damage that is caused following binding of this toxin with the unknown ligand, resulting in the loss of villous epithelial cells, and partial villous atrophy (Clarke, 2001).

(b) Heat-labile enterotoxin (LT)

LT is an oligomeric toxin that shares significant homology (ca 80% identical protein sequence) with an enterotoxin expressed by *V. cholerae*, termed cholera enterotoxin (CT) (Nataro and Kaper, 1998). In ETEC, two serotypes of LT can be

found, namely LT-I and LT-II. Both these toxins are composed of a single A subunit (28 kDa), and five larger identical B subunits (11.5 kDa), the latter of which can bind ganglioside GM1, a receptor that is distributed widely in all tissues of the body (Nataro and Kaper, 1998). It is the A subunit however that is solely responsible for the enzymatic activity of the toxin by activating adenylate cyclase (Clarke, 2001). This results in increased intracellular levels of cyclic adenosine monophosphate (cAMP), which in turn leads to diarrhoea due to an alteration in electrolyte balance, a net result of decreased sodium absorbance by villous epithelial cells and stimulated chloride secretion by crypt cells (Clarke, 2001).

1.2.2 Enteroinvasive *E. coli* (EIEC)

EIEC are an important cause of diarrhoea and are a significant cause of morbidity and mortality in young children in undeveloped countries, where sanitation and hygiene levels are of a poor standard (Robins-Browne and Hartland, 2002). EIEC strains were first shown to be capable of causing diarrhoea in volunteer studies by DuPont *et al* (1971), and have since been shown to be biochemically, genetically, and pathogenically closely related to

Shigella spp (Nataro and Kaper, 1998). Unlike ETEC, which remains within the intestinal lumen throughout the course of infection, EIEC are able to proliferate within the epithelial cells (Robins-Browne and Hartland, 2002). This process results in the inflammation and ulceration of the mucosa and ultimately leads to cell death (Clarke, 2001). This invasive phenotype, shared by both EIEC and *Shigella* spp, is dependant on the presence of a 120-140 MDa plasmid which encodes all the genes necessary for this virulence property (Clarke, 2001). Infection with EIEC is manifested clinically by fever; sever abdominal cramps, malaise, and watery diarrhoea followed by scanty dysenteric stools containing blood and mucus (Levine, 1987).

1.2.3 Enteropathogenic *E. coli* (EPEC)

As with ETEC and EIEC, EPEC are an important cause of infantile diarrhoea in the developing world (Nataro and Kaper, 1998). A number of serotypes are responsible for disease and despite not producing any classical toxins, EPEC are armed with a variety of other virulence factors (Clarke, 2001). In contrast to EIEC, EPEC do not penetrate the intestinal mucosa following infection and instead give rise to histopathological changes, termed attaching and effacing

lesions (A/E lesions) (Robins-Browne and Hartland, 2002). This histopathological phenotype is the hallmark of EPEC and other diarrheagenic *E. coli* which share this phenotype. The pathological changes include bacterial attachment and effacement of the intestinal epithelial cell microvilli (Nataro and Kaper, 1998). The intimate adherence of the bacteria to the epithelial cells causes marked changes in the host cell cytoskeleton and actin-rich pedestal formation at the site of bacterial contact (Nataro and Kaper, 1998). The genes required for the formation of A/E lesions are contained within a pathogenicity island (PAI) on the bacterial chromosome termed the locus of enterocyte effacement (LEE) (Elliott *et al.*, 1998; McDaniel *et al.*, 1995; Schmidt, 2010). Comprised of approximately 40 genes, the LEE encodes a type III secretion system (T3SS), multiple secreted proteins and a bacterial adhesin called intimin, all of which are required by EPEC for the formation of A/E lesions on enterocytes (Elliott *et al.*, 1998; Iguchi *et al.*, 2009; Jarvis *et al.*, 1995; Nataro and Kaper, 1998; Schmidt, 2010). The LEE and the TTSS it encodes are described in more detail in section 1.3.2.2.

In addition to the LEE, EPEC also possess a large virulence plasmid that includes genes encoding bundle forming pili (*bfp*) (Clarke, 2001). These Bfp are

considered to be responsible for the initial contact of EPEC with the intestine, and volunteer studies have shown that mutations in the genes required for *bfp* formation result in significantly milder illness, reinforcing their role in EPEC virulence (Robins-Browne and Hartland, 2002). Clinically, EPEC illness is characterised by fever, malaise, vomiting, and diarrhoea with large amounts of mucus in the stool. In infants, EPEC infection tends to be clinically more severe than other diarrhoeal infections, and diarrhoeal episodes can persist for as long as two weeks (Levine, 1987). The exact mechanism(s) of diarrhoea in EPEC infection however remain(s) to be fully elucidated, although several have been proposed. These include malabsorption due to the dramatic loss of microvilli in the A/E lesion, alteration of ion transport in epithelial cells by EPEC, and diarrhoea due to inflammation and increased intestinal permeability (Nataro and Kaper, 1998).

1.2.4 Enteroaggregative *E. coli* (EAEC)

EAEC, like the other pathotypes of ETEC, EPEC and EIEC, are a significant cause of diarrhoea in developing countries and are associated with acute and persistent diarrhoea, particularly in children (Clarke, 2001). EAEC derive their

name from their distinctive adherence patterns on cultured cells, but in contrast to EPEC, do not induce A/E lesions (Clarke, 2001). Instead, EAEC form a characteristic bacterial layer, widely described as a stacked-brick configuration, and experimental studies have shown that infection by EAEC results in increased mucus secretion at the mucosal surface, and characteristic mucus biofilm formation (Nataro and Kaper, 1998). Although the pathogenesis of EAEC infection is not as well understood as for some of the other diarrheagenic *E. coli* classes, several candidate virulence factors have been described. One of these virulence factors is a 38 amino acid protein known as enteroaggregative heat-stable enterotoxin (EAST1) that is functionally similar to ST (Clarke, 2001; Nataro and Kaper, 1998). More recently, a 108 kDa cytotoxin has been described although the role of this in EAEC pathogenesis remains to be fully established (Nataro and Kaper, 1998). In addition to these toxins, two fimbriae have been characterised from EAEC, termed aggregative adherence fimbriae I and II (AAFI and AAFII) (Clarke, 2001). These plasmid encoded fimbriae are believed to be responsible for the aggregative adherence (AA) phenotype of EAEC, although other as yet unknown adherence factors may play a role in the binding of the bacterium to the mucosa and in biofilm formation (Nataro and Kaper, 1998). Although the site of EAEC infection in the human intestine has yet to be

demonstrated and the clinical features of EAEC diarrhoea are not well defined, evidence suggests that either watery or mucoid diarrhoea, with little or no vomiting, largely represents the symptoms of EAEC infection (Nataro and Kaper, 1998).

1.2.5 Enterohaemorrhagic *E. coli* (EHEC)

1.2.5.1 Nomenclature – EHEC, STEC and VTEC

EHEC constitute a subset of serotypes of Shiga toxin producing *E. coli* (STEC), formerly known as Shiga-like-toxin producing *E. coli* (SLTEC), that have been associated with haemorrhagic colitis (HC) and haemolytic uraemic syndrome (HUS) in industrialised countries (Caprioli *et al.*, 2005; Chase-Topping *et al.*, 2008). Regarded as an emerging pathogen, it was initial observations by both Riley *et al* (1983) and Karmali *et al* (1985) that were responsible for the association of HC and HUS with infection by EHEC.

It was this discovery of EHEC by a number of different researchers that lead to the current nomenclature system for this diarrheagenic pathogen. Confusingly, EHEC strains are also known as verotoxigenic *E. coli* or Verocytotoxin

producing *E. coli* (VTEC), a consequence of the observations made by Konowalchuck *et al* (1977); that these strains are cytotoxic for Vero cells (Nataro and Kaper, 1998). STEC and VTEC both refer to *E. coli* strains that produce one or more toxins of the shiga toxin (Stx) family, and these terms are used interchangeably by the scientific community. Not all strains of STEC/VTEC are able to cause HC and HUS however and so the term enterohaemorrhagic *E. coli* was chosen to denote those strains that cause HC and HUS, express Stx, and carry other virulence determinants such as the LEE pathogenicity island and a large virulence plasmid (Robins-Browne and Hartland, 2002). In effect, all strains referred to as EHEC have clinical connotations, and are highly pathogenic, as opposed to many STEC/VTEC strains that are apparently of lesser virulence (Nataro and Kaper, 1998).

1.2.5.2 Clinical importance of EHEC

In humans, EHEC infection is associated with a spectrum of clinical symptoms that range from asymptomatic infection to mild-non bloody diarrhoea, HC (bloody diarrhoea), HUS, thrombocytopenic purpura and death (Boyce *et al.*, 1995). HC is a distinct clinical syndrome that presents typically with severe

abdominal cramps, bloody stools, little or no fever, and evidence of colonic mucosal oedema, erosion or haemorrhage (Griffin and Tauxe, 1991). The disease is distinguished from inflammatory colitis by the lack of significant fever and absence of an inflammatory exudate in stools (Griffin and Tauxe, 1991). For patients suffering from HC, the quantity of blood presented in faeces can vary enormously from a few streaks, to a stool comprised almost entirely of blood. In severe cases of HC this can be life threatening, particularly in the young and elderly (Coia, 1998).

Haemolytic uraemic syndrome is the most important complication of EHEC infection and this clinical syndrome is characterised by microangiopathic haemolytic anaemia, thrombocytopenia and acute renal failure. It occurs in approximately 7% of cases 5-10 days after the initial presentation of symptoms (Coia, 1998). On the basis of epidemiological, clinical and laboratory data, two major subgroups of HUS syndrome have been identified: typical (or epidemic) and atypical (or sporadic) (Su and Brandt, 1995). As is the case with HC, infants and elderly are more likely to develop HUS and mortality rates range from 5% to 10%, with a proportion of patients developing long-term sequelae (Su and Brandt, 1995).

1.2.5.3 Epidemiology of EHEC O157:H7

EHEC may belong to a number of O serogroups, but those of serogroup O157 are the most important in human disease (Clarke, 2001). *E. coli* O157:H7 was first recognised as a cause of human illness in 1982, where outbreaks of HC were associated with the consumption of undercooked hamburgers, at a chain of fast food restaurants in Michigan and Oregon, USA (Robins-Browne and Hartland, 2002). The causative organisms were found to be *E. coli* O157:H7 that were isolated from the stools of those persons affected, and from a sample of the implicated burgers (Su and Brandt, 1995). Increasing numbers of diseases relating to O157:H7 have been reported since these two outbreaks, most of which have been sporadic, although a number of community wide outbreaks in nursing homes, schools, day care centers and restaurants have also occurred (Su and Brandt, 1995).

1.2.5.4 Reservoirs of EHEC

1.2.5.4.1 Cattle

Ruminants, in particular cattle, are considered to be the most important source of human infection with *E. coli* O157:H7, being asymptomatic excretors of the

organism, which is as a transient member of their normal gut flora (Caprioli *et al.*, 2005). Prevalence studies in Scotland have shown that 23% of herds and 7.9% of cattle shed *E. coli* O157:H7 in their faeces, presenting a considerable environmental risk to humans, particularly by those animals regarded as high shedders ($>10^5$ /g) (Strachan *et al.*, 2006). Studies have shown that *E. coli* O157:H7 may remain viable in animal faeces for in excess of 20 months, and contamination of soil with effluents from agriculture, sewage and abattoirs all potentiate the risk of human infection (Beutin, 2006). Differences in prevalence and shedding of O157:H7 have also been found between countries, regions, farms and herds. In Scotland, it has been calculated that 80% of transmission arises from 20% of the most infective individual animals, termed 'super shedders' (Matthews *et al.*, 2006). Studies conducted in the United States, England and Wales all suggest that faecal shedding is more prevalent in young calves immediately after weaning, and these results have also been reproduced in experimental infection studies, where it was shown that EHEC O157:H7 shedding is longer and more intense in calves than in adult cattle (Caprioli *et al.*, 2005). Interestingly, excretion of EHEC O157:H7 by cattle has been shown to be higher during summer and early autumn, which also coincides with the seasonal peak of human infections (Karch *et al.*, 2005).

1.2.5.4.2 Non ruminants

E. coli O157:H7 has also been sporadically isolated from non-ruminant animals including: pigs, cats, dogs, chickens, horses and gulls (Nataro and Kaper, 1998).

The roles that these animals play in terms of the harbouring and transmission of *E. coli* O157:H7 remains to be fully established (Caprioli *et al.*, 2005).

1.2.5.5 Transmission of EHEC

1.2.5.5.1 Food-borne transmission

EHEC can be transmitted by a number of different routes but perhaps the most important is food-borne-transmission. Indeed, many of the largest outbreaks of *E. coli* O157:H7 have been linked to the consumption of under cooked meat, in particular, hamburgers (Karmali, 1989). *E. coli* O157:H7 can enter the food chain in a variety of ways but contamination of meat from bovine faeces during slaughter and meat processing is one of the major routes by which this happens (Karch *et al.*, 2005). Sources other than undercooked meat have also been implicated in the transmission of *E. coli* O157:H7, with unpasturised milk being one of the most commonest means of transmission (Karch *et al.*, 2005). The spectrum of vehicles implicated in the transmission of *E. coli* O157:H7 now

includes many sources far beyond the original hamburger associated outbreaks (Nataro and Kaper, 1998), with infection by *E. coli* O157:H7 even being linked to properly treated meat and dairy products, such as yoghurt and pasteurised milk (Karch *et al.*, 2005). Additionally, *E. coli* O157:H7 has also been linked to unpasturised apple juice and fermented salami. These two vehicles demonstrate the potential acid tolerance of *E. coli* O157:H7, which is able to grow and survive in extremely acidic conditions that would normally destroy other pathogens (Nataro and Kaper, 1998).

More recently, raw vegetables and fruits have also become major vehicles for the transmission of *E. coli* O157:H7, probably through contamination with cattle faeces during harvesting and processing (Karch *et al.*, 2005). Lettuce in particular has been implicated as a vehicle in several outbreaks (Nataro and Kaper, 1998) and uncooked radish sprouts were responsible for an outbreak of *E. coli* O157:H7 infection in Sakai City, Japan, where more than 6000 school children were affected (Beutin, 2006).

1.2.5.5.2 Water-borne and environmental transmission

Non food-borne vehicles have also been implicated in the transmission of *E. coli* O157:H7. Drinking of contaminated water from sources such as recreational water, wells and municipal systems, have all been associated with outbreaks (Nataro and Kaper, 1998). Research has shown that *E. coli* O157:H7 is able to survive in both drinking water and recreational water for many weeks, and evidence suggests that *E. coli* O157:H7 can enter a viable but non-cultural state after persisting in water for 12 weeks (Henderson *et al.*, 1999). Along with the contamination of preserved foods and vegetables, water contamination is of particular concern in terms of the scale of risk that it poses to public health (Nataro and Kaper, 1998).

An increasing number of outbreaks of *E. coli* O157:H7 infection have also occurred amongst people visiting music festivals, fairs, parties and petting zoos, particularly in the summer months (Caprioli *et al.*, 2005). The sources of infection are believed to be wells, exposure to mud contaminated with cattle faeces, and direct contact with animals (Caprioli *et al.*, 2005). Transmission through direct animal contact has been reported in outbreaks and sporadic

infections by *E. coli* O157:H7, and visiting farms or petting zoo and direct contact with animal faeces is now considered a significant hazard (Karch *et al.*, 2005).

1.2.5.5.3 Person-to-person transmission

The infective dose for *E. coli* O157:H7 is very low and has been estimated to be in the order of 100 to 200 organisms (Nataro and Kaper, 1998). Because of this low infectious dose, person-to-person faecal-oral transmission of *E. coli* O157:H7 has become one of the major routes of infection in daycare settings and in any institution where poor hygiene and close contacts may exist (Karch *et al.*, 2005). In an outbreak of *E. coli* O157:H7 among kindergarten children, who probably acquired their infection through drinking raw milk at a farm, the occurrence of secondary cases among family members suggested that person-person transmission was responsible for the spread of infection (Karmali, 1989). This case reinforces the importance of person-to-person transmission as a vehicle for *E. coli* O157:H7 infection, and highlights the importance of good hygiene and sanitation as a control measure against the potential spread of this pathogen.

1.2.5.5.4 Incidence of EHEC

To date, the largest outbreaks of EHEC infection have affected thousands of individuals. However, sporadic EHEC infections make up the vast majority of the disease burden for this organism, and evidence suggests that the frequency of sporadic cases of EHEC infection is on the increase (Nataro and Kaper, 1998). Like outbreaks, sporadic infections due to *E. coli* O157:H7 have become a major problem in both the United States and Canada, where geographical differences in the incidence of this pathogen can be seen (Nataro and Kaper, 1998). In the USA, sporadic cases of *E. coli* O157:H7 and HUC are more common in the northwestern states than the southern states; and in Canada, infection by *E. coli* O157:H7 has been reported more frequently from western provinces than eastern provinces (Griffin and Tauxe, 1991). In the USA, this higher incidence of *E. coli* O157:H7 infection in the north may be explained by larger rural populations and increased contact with farm animals (Beutin, 2006). The annual incidence of reported *E. coli* O157:H7 infections in Canada and the United States ranges from 1.7 to 5.3 per 100,000 persons, and within the United States alone, it has been estimated that approximately 73,000 cases of *E. coli* O157:H7 infection occur annually (Steele *et al.*, 2007).

E. coli O157:H7 is also an important pathogen in Europe and Japan, and has been isolated in many other parts of the world including Asia, Africa and South America (Boyce *et al.*, 1995). In the developed countries of the northern hemisphere, *E. coli* O157:H7 follows a distinct seasonality of infection with most sporadic cases being reported in the warm summer months (Nataro and Kaper, 1998). This seasonal variation may reflect the ecology of *E. coli* O157:H7 or may be due to some unknown factor (s) (Boyce *et al.*, 1995). The reasons for the increase in sporadic *E. coli* O157:H7 infections can be explained in part by the establishment of more sophisticated EHEC detection and isolation methods, and to increased surveillance, resulting in more identified clinical cases (Beutin, 2006). Other factors which may explain the increase in frequency of sporadic *E. coli* O157:H7 infections include the industrialisation of agriculture coupled with mass production of livestock and animals; and behavioral changes in consumer habits including the increased consumption of raw vegetables, dairy products, meat products and undercooked meat (Beutin, 2006).

1.3 Pathogenesis and virulence of EHEC O157:H7

1.3.1 Shiga toxin

One of the most important and defining virulence characteristics of *E. coli* O157:H7 is its ability to produce one or more Shiga toxins. Shiga toxins, namely Stx1 and Stx2, occur in two major antigenic groups and have an A-B structure (Robins-Browne and Hartland, 2002). Stx1 has been shown to be indistinguishable from the toxin produced by *Shigella dysenteriae* type 1, whilst Stx2 exhibits far more diversity, sharing only 56% amino acid homology with Stx1 (Mead and Griffin, 1998). Variants of Stx2 have been identified, including Stx2c, d, e, f and g. Human disease however is most significantly correlated with Stx2 and Stx2c (Spears *et al.*, 2006). The structural genes encoding Stx1 and Stx2 are carried on a bacteriophage and are considered to be transferred to the chromosome of *E. coli* O157:H7 and other host EHEC following phage lysogeny (Osawa *et al.*, 2000).

1.3.1.1 Mode of action of Shiga toxin

Across all members of the Shiga toxin family, the basic A-B subunit structure is conserved. A single 32 kDa A subunit is proteolitically cleaved to yield a ca. 28

kDa peptide (A₁) and a 4 kDa peptide (A₂) linked by a disulphide bond (Nataro and Kaper, 1998). The A₁ peptide is solely responsible for the enzymatic activity of the toxin and is bound to a pentamer of five identical 7.7 kDa B subunits via the A₂ peptide of the A subunit (Nataro and Kaper, 1998). The eukaryotic cell surface receptor for members of the shiga toxin family is globotriaosylceramide (Gb₃), and it is this glycolipid receptor that is bound by the B subunit of Shiga toxin (Paton and Paton, 1998). Following initial binding by the B subunit to a target cell membrane, internalisation of toxin is thought to occur through a process of receptor-mediated endocytosis (Paton and Paton, 1998). The toxin is endocytosed in clathrin-coated pits and processed by retrograde transport through the Golgi apparatus and endoplasmic reticulum, where the A₁ subunit exerts its N-glycosidase activity following translocation to the cytoplasm (Nataro and Kaper, 1998). This leads to the cleavage of 28S ribosomal RNA which disrupts the integrity of the ribosome and inhibits protein synthesis, ultimately leading to cell death (Robins-Browne and Hartland, 2002; Spears *et al.*, 2006).

1.3.1.2 Role of Shiga toxin in intestinal disease

The involvement of Stx in intestinal disease has been supported by a large number of studies involving animal models, cell lines and genetically mutated strains of pathogenic bacteria (Nataro and Kaper, 1998). Studies in rabbits for example have shown that Stx has direct enterotoxic properties on the rabbit ileum, where Gb₃ containing absorptive villus epithelial cells are targeted (Paton and Paton, 1998). Work by Fontaine *et al* (1988) demonstrated that in monkeys, *S. dysenteriae* type I causes less severe disease when it has a specific mutation for Stx, as compared to Stx positive *S. dysenteriae* type I in feed challenge studies (Nataro and Kaper, 1998). Although these studies strongly support an association between Stx and intestinal disease in animals, the mechanisms involved in Stx pathogenesis in humans are not well understood. Direct exposure of enterocytes to Stx may explain diarrhoeal disease in humans although the presence of Gb₃ on human enterocytes has yet to be demonstrated (Paton and Paton, 1998). Gastrointestinal pathology as a result of systemic Stx exposure has also been suggested, although the exact mechanisms of toxin translocation from the site of colonisation in the gut lumen, to the underlying tissues and blood stream remains to be fully elucidated (Paton and Paton, 1998). Studies using cell lines have shown however that Stx can translocate across

epithelial cells without obvious cellular disruption, presumably through a transcellular pathway; and this may in part explain the systemic sequelae observed following infection with Stx producing *E. coli* O157:H7 (Nataro and Kaper, 1998).

1.3.1.3 Role of shiga toxin in HUS

HUS is thought to develop when Stx produced in the intestine enters the blood and binds to Gb₃ rich endothelial cells in the kidneys, although toxin has never been detected in the blood of HUS patients (Mead and Griffin, 1998; Nataro and Kaper, 1998). In addition to the direct action of Stx on renal endothelial cells, clinical studies have strongly associated both the severity and outcome of HUS with the production of pro-inflammatory cytokines (Nataro and Kaper, 1998). For example, interleukin 6 (IL-6) has been shown to be elevated in both the serum and urine of patients suffering from HUS, suggesting that renal pathology in HUS is a consequence of both the host immunological response to Stx; and the direct cytotoxic action of Stx (Nataro and Kaper, 1998). The typical features of renal pathology in HUS includes swollen and detached glomerular endothelial cells; deposition of fibrin and platelets within glomeruli and other

parts of the renal microvasculature; and capillary occlusion resulting in reduced blood flow to the kidneys (Paton and Paton, 1998). Other pathological abnormalities found in HUS include decreased platelet counts; increased plasma platelet aggregation activity; the occurrence of abnormal factor VIII; and microangiopathic haemolytic anaemia (Karmali, 1989).

1.3.2 Attaching and effacing (A/E) lesions

E. coli O157:H7 along with other EHEC, EPEC and the mouse pathogen *Citrobacter rodentium* (CR), all belong to a family of enteric bacteria that form A/E lesions and cause diarrhoea (Deng *et al.*, 2004). As described earlier in relation to EPEC pathogenesis, A/E lesions are characterised by the effacement of microvilli on the surface of intestinal epithelial cells, an intimate adherence of bacteria to the host cell membrane and the formation of organised cytoskeletal structures called actin-pedestals underneath the host cell membrane at the site of bacterial attachment (Campellone and Leong, 2003; Kaper *et al.*, 1998). Studies have shown that both EHEC and EPEC have the ability to form actin pedestals on cultured cells and this feature correlates with the ability to colonise the intestine and cause disease in human and animal hosts (Campellone and

Leong, 2003). Intriguingly, LEE-negative EHEC isolates are regularly isolated from patients with severe disease. Although unable to form A/E lesions, these strains are invasive on cells *in vitro* where they are internalized within a membrane-bound vacuole; suggesting that aside from the LEE PAI that confers the A/E lesions, other unidentified factors can enhance the virulence potential of these strains (Girardeau *et al.*, 2009; Luck *et al.*, 2005).

1.3.2.1 Role of LEE in A/E lesion formation

For EHEC O157:H7, A/E lesion formation and pathology is determined by the LEE which contains most of the genes critical for lesion formation (Deng *et al.*, 2001; Perna *et al.*, 1998). It has been demonstrated that when cloned onto a plasmid, the LEE from EPEC strain E2348/69 can confer an A/E phenotype in the non pathogenic *E. coli* strain K12 (McDaniel and Kaper, 1997). In contrast, when the LEE of EHEC was cloned into the same K12 strain, it was unable to form A/E lesions (Elliott *et al.*, 2000). Publication of the complete LEE sequence from three A/E pathogens; EPEC O127:H6 strain E2348/69; EHEC O157:H7 strain EDL933 and rabbit O15:H- strain RDEC-1 have shown that the LEE is highly conserved, suggesting a common origin for all three pathogens (Deng *et al.*, 2001). It has

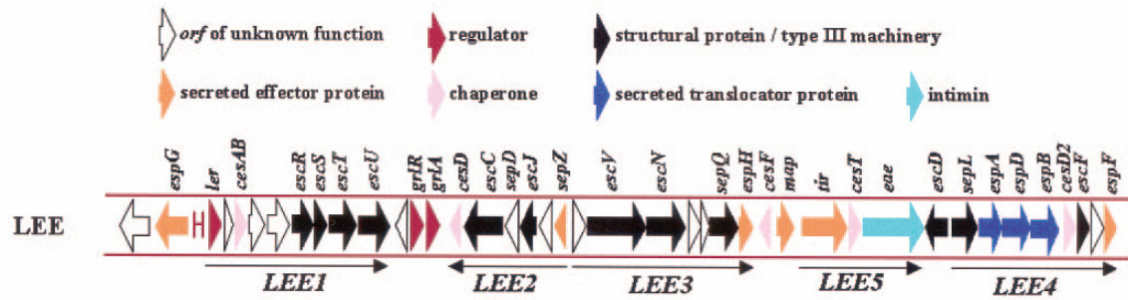
also been shown that the LEE has a considerably lower G+C content (38%) as compared to the average *E. coli* genome (51%), indicating that the LEE was probably acquired by these pathogens following horizontal gene transfer from another species (Wieler *et al.*, 1997). Exactly where the LEE originated and when it was transferred however still remains to be fully established (Deng *et al.*, 2001).

1.3.2.2 LEE and type III secretion (T3S)

The discovery of T3S was first made in the 1980's and 1990's, when researchers studying *Yersinia* found that the bacteria produced proteins called Yops (*Yersinia* outer proteins). Thought to be associated with the outer membrane, it was soon realised that these proteins lacked classical signal sequences and were not secreted via the general secretion-dependent pathway (type II secretion) (Gauthier *et al.*, 2003; Michiels *et al.*, 1990). A new secretion system was therefore proposed and this later became known as a Type III secretion system (T3SS) (Gauthier *et al.*, 2003).

EHEC and EPEC along with many other pathogenic Gram negative bacteria, including those that infect plants, utilise a T3SS to secrete and inject bacterial effector proteins into the cytosol of host cells (Hueck, 1998). These bacterial effector proteins, which often resemble eukaryotic factors, are able to subvert host cell signaling pathways, and are also responsible for the cytoskeletal rearrangements that occur during A/E lesion formation following infection by EHEC and other A/E pathogens (Hueck, 1998). In general, the T3SS comprises a basal apparatus, with proteins present in the inner and outer bacterial membranes, and a needle complex known as a translocon which allows the injection of effector proteins through the host cell membrane (Figure 1.1B) (Germendia *et al* 2005). In EHEC O157:H7 strain EDL933, the LEE is 43.359 kb in size and is composed of 54 ORF's, of which 13 fall within a putative prophage region known as 933L (Figure 1.1A) (Perna *et al.*, 1998). The LEE encoded genes are organised into five major operons (*LEE1* through *LEE5*) and it is *LEE1*, *LEE2* and *LEE3* operons that contain the type III basal apparatus genes (Figure 1.1A-B). The *LEE4* operon encodes EscF and the *E. coli*-secreted proteins A, B and D (EspA,-B,-D) which are involved in *E. coli* O157:H7 needle and translocon complex formation, as well as factors whose function remains unclear (Figure 1.1B). The *LEE5* operon encodes an outer membrane adhesin called intimin, its

A



B

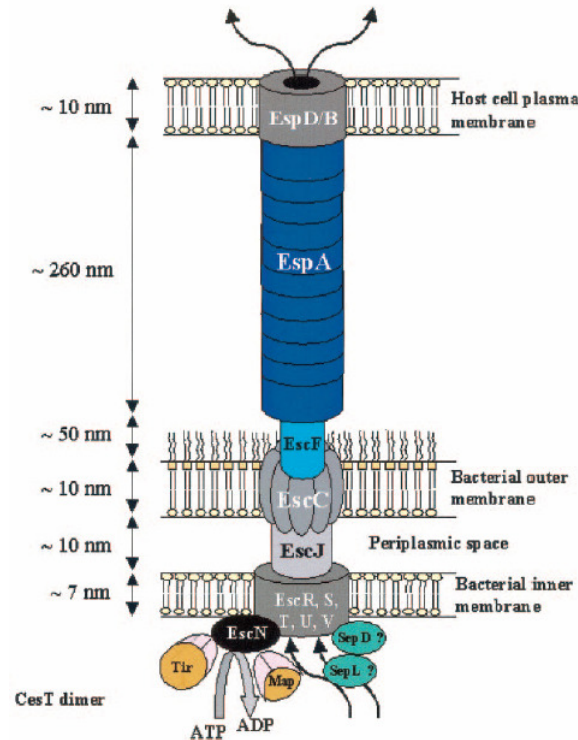


Figure 1.1 Locus of enterocyte effacement (LEE) encoding a type III secretion system (T3SS) (A) Genetic organisation of the LEE PAI in EHEC O157:H7. (B) Translocon apparatus of the LEE encoded T3SS. The basal body of the TTSS is composed of the secretin EscC, the outer membrane proteins EscR-EscV and the lipoprotein EscJ which connects the inner and outer membrane ring structures. EscF forms the needle structure where EspA polymerises to form the EspA filament. EspB and EspD form the translocation pore in the host cell membrane, providing a continuous but gated channel from the bacteria to the eukaryotic cell. EscN provides energy to the system by hydrolysing ATP to ADP. Both images were reproduced from Garmendia *et al* (2005).

translocated receptor Tir, and a Tir chaperone known as CesT (Figure 1.1A-B) (Garmendia *et al.*, 2005; Gauthier *et al.*, 2003). Following the assembly of the type III basal apparatus and EspA needle complex, EspB and Tir are translocated to the host epithelial cell where Tir is integrated into the plasma membrane. It is the binding of the translocated Tir with intimin on the bacterial surface that is responsible for the formation of intimate attachment during infection with *E. coli* O157:H7 (Campellone and Leong, 2003; Gauthier *et al.*, 2003). Other LEE encoded effectors (Map, EspF, EspG, and EspH) have been identified which are involved in the modulation of the host cytoskeleton, although these effectors are dispensable for A/E lesion formation (Deng *et al.*, 2004; Spears *et al.*, 2006). In addition to binding Tir, intimin has also been shown to bind to host derived receptors (Spears *et al.*, 2006). It has been suggested that the interaction of different antigenic subtypes of intimin with these endogenous host receptors may be responsible for the differences in tissue tropism observed for both EHEC and EPEC (Spears *et al.*, 2006). Sequence comparison analyses has shown that while most of the genes coding for the T3SS show greater than 95% identity among the two sequenced strains from EPEC and EHEC, strains E2348/69 and EDL933 respectively (Deng *et al.*, 2001), those coding the secreted proteins EspA, EspB, EspD and EspE are far more diverse (84.63, 74.01, 80.36,

and 66.48% homology respectively) (Kresse *et al.*, 1999). Somewhat surprisingly given its role in A/E lesion formation, Tir is the most divergent of all the Esps, sharing only 58% homology between EPEC and EHEC (Campellone *et al.*, 2004). This divergence of Tir is illustrated by the fact that unlike EPEC, EHEC does not accumulate tyrosine phosphorylated proteins beneath adherent bacteria, and requires other bacterial factors in addition to Tir for the formation of actin pedestals (Campellone and Leong, 2003; DeVinney *et al.*, 1999). Although the LEE PAI of EHEC and EPEC is highly conserved, differences in pedestal generation highlight just one of the many important variations that exist between these two A/E pathogens.

1.3.2.3 Regulation of the LEE and T3S

Genetic determinants governing bacterial virulence are tightly regulated so as to ensure their appropriate expression when environmental conditions are at their optimum; avoiding unnecessary wastage of energy and risking premature stimulation of host immune responses that may otherwise compromise the infectious process. In *E. coli* O157:H7, LEE gene expression is regulated in an extremely complex manner governed by environmental signals; by regulatory

proteins encoded within the LEE PAI itself; by a system of global regulators found in *E. coli*; and a variety of regulators that have been acquired via horizontal gene transfer. A summary diagram depicting some of the better characterised transcriptional regulators of the LEE PAI and where they feed in to the LEE operon are shown in Figure 1.2.

1.3.2.4 Intrinsic regulation of the LEE

LEE gene expression is under the direct control of the LEE-encoded regulator (*Ler*), which is the first gene in the *LEE1* operon (Hacker and Kaper, 2000; Sperandio *et al.*, 2002). *Ler* proteins (predicted to be 15.1 kDa) of A/E pathogens are highly homologous in structure and function (Mellies *et al.*, 2007) and share sequence similarity to H-NS, a global regulator in *E. coli* that can silence the expression of foreign DNA; including the LEE PAI. These two regulators share most sequence similarity at the C-terminus of H-NS (predicted to be a DNA binding domain) (Bustamante *et al.*, 2001; Mellies *et al.*, 2008). *Ler* is essential for A/E lesion formation as it activates the transcription of those operons downstream of *LEE1*, *LEE2-5*, as well as other operons (e.g. *grlRA*, *espG* and *map*) and genes within the LEE; in a cascade fashion. *Ler* has also been shown to

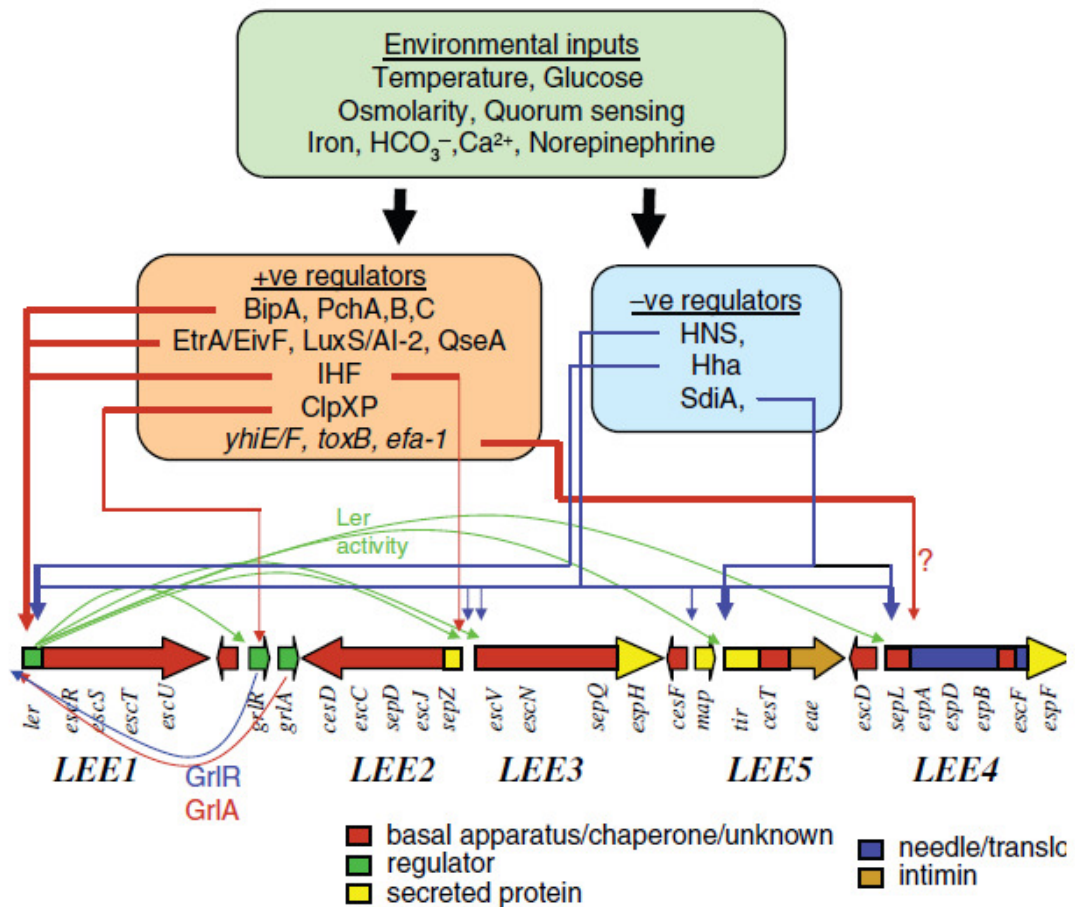


Figure 1.2 Schematic summary of the complex regulation of the LEE PAI in EHEC O157:H7. The diagram shows T3S is regulated at multiple levels and by many factors encoded both within and out with the LEE. Transcriptional regulators are annotated as shown along with their input site into the LEE operon. The mechanisms of action of individual regulators are discussed in more detail in the main text. This image was reproduced from Spears *et al* (2006).

regulate the expression of genes and genetic elements (prophage and plasmid encoded) encoded outside of the LEE, for example *stcE* (plasmid O157 (pO157) encoded zinc metalloprotease that is secreted by the *etp* type II secretion system) and *lpf* (long polar fimbriae) (Mellies *et al.*, 2007). Ler is therefore regarded as a global regulator able to coordinate LEE and virulence gene expression in EHEC.

In EPEC, it has been shown that Ler and H-NS can occupy the same binding sites upstream of the *LEE2* and *LEE5* promoters (Mellies *et al.*, 2007). As Ler is considered to have a lower dissociation constant (K_d) than H-NS, these findings have led to the suggestion that Ler disrupts H-NS from binding to the upstream and downstream regions of target genes and operons (Berdichevsky *et al.*, 2005; Dame *et al.*, 2005; Mellies *et al.*, 2008); where biochemical analyses have shown that H-NS silences transcription by binding over extended regions of DNA, with the capacity to bridge and loop DNA (Dame *et al.*, 2005; Dorman 2004; Dorman and Kane, 2009). Ler does not activate the transcription of all H-NS repressed genes (e.g *proU*) and as such, it is not considered as a general H-NS antagonist, but rather as a specific global activator of virulence gene expression (Mellies *et al.*, 2007). It has also been reported that Ler has the capacity to auto-regulate its own expression under certain environmental conditions; indicating that local

Ler concentrations at the *LEE1* promoter may be important for the proper regulation of LEE gene expression in both EPEC and EHEC (Berdichevsky *et al.*, 2005; Mellies *et al.*, 2008; Mellies *et al.*, 2011).

The *LEE1* regulatory region located immediately upstream of *ler* is crucial in controlling expression of the LEE. In EHEC, this region has been reported to contain two promoters, P1 and P2 that are respectively distal and proximal to the *ler* translation start site (Sperandio *et al.*, 2002). The global regulator LEE activator and repressor proteins, GrlA and GrlR respectively, are located between *LEE1* and *LEE2*; and these proteins have the ability to control *LEE1* transcription in a positive feedback loop with Ler (Barba *et al.*, 2005; Iyoda *et al.*, 2006). Studies by Barba *et al.* (2005) have shown the expression of LEE-encoded proteins to be significantly reduced following mutation of *grlA*; and increased following mutation of *grlR*. It is suggested that since *grlA* lies immediately downstream of *grlR*, these two genes form a bicistronic operon; where GrlR may directly repress GrlA by protein-protein interactions that result in the down regulation of *ler* expression (Iyoda *et al.*, 2006; Jobichen *et al.*, 2007).

Recent works by Islam *et al* and Bustamante *et al* (2011) have also indicated that GrlA may have dual functions at the *LEE1* regulatory region (promoter region upstream of *LEE1*); by counteracting the global repressor H-NS, and by activating the P1 promoter. These authors argue that GrlA may function as a true transcription activator by directly accelerating one of the transactions between RNA polymerase holoenzyme and the *LEE1* promoter. In addition to regulating T3S, GrlRA have also been shown to coordinate LEE expression with enterohaemolysin gene expression; and can reciprocally regulate T3S and motility gene expression in EHEC O157:H7 (Iyoda *et al.*, 2006 and Saitoh *et al.*, 2008).

1.3.2.5 Horizontally acquired regulators of the LEE

The EAF plasmid of EPEC harbours the operon *perA-C* (plasmid encoded regulator C), termed *pchA-C* in EHEC (plasmid encoded regulator C homologue), which activates *LEE1* and then via Ler, the rest of the *LEE* (Porter *et al.*, 2004; Porter *et al.*, 2005; Iyoda and Watanabe, 2004). In EPEC, PerA directly activates *perABC* and *bpf* transcription, conferring a localised adherence (LA) phenotype to EPEC on host cells (Lida *et al.*, 2010; Mellies *et al.*, 2007). PerA,

which shares sequence similarity to the AraC family of transcriptional regulators, directly activates *bfp* by binding to the promoter region upstream of *bfp*. The promoter regions upstream of *perA* and *bfp* share a significant amount of homology, and as such, *perA* is also subject to auto-regulation by PerA (Ibarra *et al.*, 2003; Mellies *et al.*, 2007). The *perB* gene product has not been well characterised, however, PerC is a known transcriptional activator of Ler (Bustamante *et al.*, 2001; Porter *et al.*, 2004).

EHEC do not harbour an EAF plasmid, however seven PerC like homologues (*pchA-E*, *pchX* and *pchY*) can be found in the EHEC chromosome; nearly all of which are prophage (OI) associated (Iyoda and Watanabe 2004; Porter *et al.*, 2005). Of these, PchA, PchB and PchC have been demonstrated to activate LEE expression through Ler; where mutation in these genes reduces T3SS translocon secretion and adherence of EHEC to tissue culture cells (Iyoda and Watanabe, 2004). These proteins also regulate genes encoded outside of the LEE on OIs (Iyoda and Watanabe 2004; Porter *et al.*, 2005), and PerC like proteins (YfdN) have also been identified on prophage elements in UPEC, *Salmonella enterica* serovar Typhimurium (ST64 phage), *Shigella flexneri* (SfV phage) and *E. coli* K12 (Mellies *et al.*, 2007).

The exact mechanism of action of these regulators is unknown although it is hypothesised that in EHEC and EPEC, PerC and PerC like proteins activate LEE transcription in the presence of other regulatory proteins, such as IHF (DNA bending protein which positively regulates LEE expression); and perhaps other factors that have yet to be identified (purified PerC has not been shown to bind to the LEE regulatory region *in vitro* (Porter *et al.*, 2005). The function (s), if any, of PchD and PchE in EHEC remain to be determined, although these proteins, which are much smaller than the PchA-C proteins, are known not to affect LEE gene expression in a K12 background (Porter *et al.*, 2005).

1.3.2.6 *E. coli* regulators of LEE

Given Ler's pivotal role in A/E lesion formation (an EHEC *ler* mutant does not form A/E lesions on tissue culture cells (Elliot *et al.*, 2000) and its role as a master regulator of the LEE, its perhaps unsurprising that several transcription factors exert their control over *LEE1* expression via Ler; either directly or indirectly. Fis and IHF for example positively regulate *LEE1* by switching on *ler* expression; subsequently Ler then activates *LEE2-5* transcription in a regulatory cascade (Bhatt *et al.*, 2011; Umansci *et al.*, 2002; Yona-Nadler *et al.*, 2003). BipA, a protein

sharing amino acid homology with the eukaryotic ribosome-binding elongation factor G, also increases LEE transcription by promoting steady state transcript levels of *ler* (Grant *et al.*, 2003; Bhatt *et al.*, 2011). Moreover, BipA is reported to be required for intimin proteolysis in EPEC (Grant *et al.*, 2003; Bhatt *et al.*, 2011). Examples of LEE repressors include the H-NS homologue, Hha, which represses *ler* transcription (Sharma and Zeurner 2004); and the acid response regulator GadX, which represses *ler* transcription indirectly via the repression of the LEE activator PerC (Mellies *et al.*, 2007). Hha has also been shown to form a complex with H-NS which has the capacity to regulate α haemolysin gene expression in response to temperature and osmolarity (Mellies *et al.*, 2007; Nieto *et al.*, 2000).

Additional repressors of the LEE can also be found in a second cryptic T3SS, termed ETT2 (*E. coli* type III secretion system 2), that is reported to be present in the majority of *E. coli* strains on prophage region OI-115 (Mellies *et al.*, 2007; Zhang *et al.*, 2004). Mutations in two ETT2 encoded genes, *eivF* and *etrA*, increase EspA and EspB translocon secretion; and over-expression of these proteins represses T3S in a high secreting O26:H- EHEC strain (Zhang *et al.*, 2004). In addition, the two component system regulators YhiEF control effector protein secretion; where they repress LEE gene expression not via *ler*, but at

LEE2 and *LEE4* (Mellies *et al.*, 2007; Tatsuno *et al.*, 2004). In EHEC, LEE gene expression is also regulated by the RcsC-B-D phosphorelay system through RcsB and the previously mentioned PchA; where RcsB represses PchA to down-regulate *ler* expression; effectively repressing the entire LEE (Tobe *et al.*, 2005). This system is positively regulated via the EHEC specific protein, GrvA, which activates LEE gene expression via *LEE1* (Tobe *et al.*, 2005).

1.3.2.7 Environmental regulation of LEE

Quorum sensing (QS) is a cell to cell signaling system in bacteria mediated through the secretion and sensing of host and bacterial hormones known as autoinducers (epinephrine, norepinephrine and autoinducer-3) (Sperandio *et al.*, 2003). These signalling molecules allow bacteria to sense their environment and gauge population density levels both within and between bacterial species which share the same environment (Sperandio *et al.*, 2003). When these compounds reach a certain threshold, they are able to interact with transcription factors to regulate gene expression; enabling the bacteria to rapidly respond to the prevailing environmental conditions (Sperandio *et al.*, 1999; Sperandio *et al.*, 2001; Sperandio *et al.*, 2003).

In EHEC, quorum sensing *E. coli* regulator A (QseA) is a member of the LysR family of transcriptional regulators that is able to directly bind to the *LEE1* regulatory region and control LEE gene expression via Ler (Sperandio *et al.*, 2002). QseA also has the capacity to bind other promoter regions and genes carried on OIs, as well as to its own promoter (Kendall *et al.*, 2011; Russell *et al.*, 2007; Sharp and Sperandio, 2007). Via a hierarchical regulatory cascade, QseA is induced through the major two-component system QseBC (Hughes *et al.*, 2009; Kendall *et al.*, 2011; Sperandio *et al.*, 2002). QseBC is known to control a number of genes involved in virulence and motility; where QseBC exerts its control over flagella expression via the master flagella regulators *FlhDC* (Kostakioti *et al.*, 2009; Sperandio *et al.*, 2002). QseC has been shown to harbour phosphatase activity and it is proposed that QseC can exert control over QseB by interfering with its dephosphorylation mechanism; allowing *flhDC* expression and the positive regulation of motility (Kostakioti *et al.*, 2009). QseB binds to high and low affinity sites upstream of *qseBC* implying it is subject to auto-regulation by QseB (Clark and Sperandio, 2005). In UPEC, in the absence of QseC ($\Delta qseC$), QseB is able to increase its own transcription, where it then goes on to repress the expression of multiple genes involved in type I pili, curli and flagella formation (Kostakioti *et al.*, 2009). This same *qseC* deletion strain was also

shown to be adversely affected for intracellular bacterial community formation. These studies support a role for QseC as a key regulator required for UPEC and EHEC virulence gene expression. Another two-component system in EHEC which has the capacity to influence virulence gene expression is QseEF. QseF activates the transcription of the EHEC effector protein EspFu which is required for actin pedestal formation; where a *qseF* deletion strain is unable to form A/E lesions (Reading *et al.*, 2007). The *qseEF* operon also encodes an OMP protein, QseG, which is believed to be necessary for Tir translocation (Reading *et al.*, 2009). As for QseA, QseEF is induced via the major QS regulator QseC.

In bacteria, the stringent response is a global regulatory system that is triggered by various nutritional and metabolic stresses; for example starvation. In *E. coli*, this system is mediated via the signalling molecule ppGpp, whose concentration increases in starved cells to control gene transcription in both a positive and negative manner (Chatterji and Ojha 2001; Nakanishi *et al.*, 2006). As ppGpp synthesis and degradation is rapid, this allows bacteria to adapt quickly to changes in their environment. In *E. coli*, ppGpp turnover is catalysed by two enzymes, relA (synthase) and SpoT (hydrolase synthase) which respond differently to environmental conditions and bacterial growth phase (Mellies *et*

al., 2007; Nakanishi *et al.*, 2006). Another protein, DksA, has also been identified as a critical component of this system and affects ppGpp mediated activation and repression of gene transcription at target promoters (Nakanishi *et al.*, 2006). Specifically in EHEC, the stringent response has been shown to enhance LEE gene expression directly through the global LEE activators Ler and Pch; where ppGpp directly activates *ler* and *pch* promoters in an *in vitro* transcription system (Nakanishi *et al.*, 2006).

The ubiquitous SOS response system in bacteria is regulated through proteins involved in DNA replication, repair and metabolism. The system is activated in response to DNA damaging agents (e.g. antibiotics, UV radiation and mitomycin C) (Janion, 2008). In EHEC, Stx toxin production is co-regulated with the induction of the lysogenic Stx-encoding bacteriophage which is also linked with the induction of the bacterial SOS system (Matsushiro *et al.*, 1999; Muhldorfer *et al.*, 1996; Plunkett *et al.*, 1999). *stx* genes are encoded in the late lytic region of bacteriophages, interestingly with a number of EHEC effector proteins. As such, the SOS response has been postulated to coordinate DNA damage with phage lysis, Stx production and effector protein secretion (Croxen and Finlay, 2010; Fuchs *et al.*, 1999; Tree *et al.*, 2009). In EPEC, mitomycin C

increases the expression of the *LEE2* and *LEE3* operons, as well as the non-LEE encoded effector A (*nleA*) that is encoded on OI-71 but secreted via the T3SS (Mellies *et al.*, 2007). This regulation is mediated by LexA and RecA, two major players in the bacterial SOS response. LexA is a transcriptional regulator which binds to an SOS box present in some operon promoter regions, such as found upstream of the *LEE2* and *LEE3* promoters on the LEE PAI, repressing transcription (Mellies *et al.*, 2007). RecA counteracts LexA mediated repression by catalytically cleaving and inactivating LexA; so promoting transcription (Mellies *et al.*, 2007).

1.3.2.8 Post-transcriptional and post-translational control of LEE

Recent studies have shed light on the increasing importance of post-transcriptional and post-translational control of LEE gene expression. CsrA (carbon storage regulator A) for example is a small RNA binding protein that regulates flagella and T3S in EHEC (Bhatt *et al.*, 2011; Wei *et al.*, 2001); and represses glycogen biosynthesis and biofilm formation in non-pathogenic *E. coli* (Wang *et al.*, 2005). CsrA is highly conserved among different species of bacteria, where it governs pleiotrophic phenotypes associated with cellular

metabolism and virulence (Romeo, 1998; Timmerman and Melderer, 2010). In EPEC, CsrA is known to bind to, and is believed to stabilise, the *LEE4* transcript (Bhatt *et al.*, 2011). CsrA is also known to promote A/E lesion formation by activation the T3SS inner membrane protein EscD; through an intermediate regulator (Bhatt *et al.*, 2011; Timmerman and Melderer, 2010). At increased concentrations, CsrA has a dual function, when it is believed to repress LEE expression by reducing the transcript stability of the global LEE regulators, *grlRA* (Bhatt *et al.*, 2011).

Another important post-transcriptional regulator of the LEE in EHEC and EPEC is the small RNA chaperone, *hfq* (Bhatt *et al.*, 2011). As with CsrA, Hfq has the capacity to influence a large number of cellular processes (inactivation of *hfq* in *E. coli* causes cells to become extremely sensitive to environmental stresses (Bhatt *et al.*, 2011; Tsui *et al.*, 1994) as well as being a virulence regulator. In EHEC, Hfq is believed to repress *LEE* transcription via two mechanisms that depend on the growth phase of the cell. During exponential growth, Hfq has the capacity to reduce GrlA expression by de-stabilising *grlRA* transcripts. This in turn reduces levels of Ler, and as such, Hfq is able to silence the whole of the LEE PAI (Bhatt *et al.*, 2011; Hansen and Kaper, 2009). During stationary phase

growth, Hfq mediated repression of the LEE is independent of *grlRA*, where this time *ler* translation is adversely affected (Bhatt *et al.*, 2011; Hansen and Kaper, 2009; Shakhnovich *et al.*, 2009). In addition to the non-catalytic RNA-binding protein Hfq, RNase E (single strand specific endoribonuclease E) exerts post-transcriptional control over LEE expression (Lodato and Kaper, 2009). RNase E is a component of the bacterial degradosome that is involved in the maturation and degradation of RNA species, and is believed to antagonistically co-regulate several transcripts with Hfq; where it is responsible for the post-transcriptional processing of the *LEE4*-encoded *sepLspADB* transcript in EHEC (Bhatt *et al.*, 2011; Lodato and Kaper, 2009).

ClpXP is a post-translational regulator of the LEE that has been extensively characterised (Bhatt *et al.*, 2011; Iyoda *et al.*, 2005; Tomoyasu *et al.*, 2005). In EHEC, ClpX and ClpP degrade damaged and incomplete proteins and proteins involved in various stress responses. These proteins are also believed to influence LEE gene expression by degrading GrlR; so allowing GrlA to activate *ler* and the LEE regulatory cascade (Iyoda *et al.*, 2005). In EHEC, ClpXP is believed to activate *ler* in an RpoS and GrlR dependant manner, where ClpXP is believed to degrade RpoS. In addition, deletion of *grlR* circumvents the need for

functional *grlR*; and in a *grlR* mutant, ClpXP levels and stability are elevated; suggesting that *grlR* may be a direct substrate for this protease (Bhatt *et al.*, 2011; Iyoda *et al.*, 2005). Interestingly, it has recently been shown that ClpXP has the ability to regulate flagella gene expression through *flhDC* transcriptional and post-translational control mechanisms. This repression is also believed to be mediated via *grlRA* (Kitagawa *et al.*, 2011). Small RNA mediated control of LEE gene expression has also been observed in *E. coli*. For example, DsrA is as an 87 nucleotide un-translated transcript that modulates gene expression by antisense pairing with its target mRNAs in the presence of Hfq (Bhatt *et al.*, 2011; Lease *et al.*, 2000). When over-expressed, DsrA activates *ler* in an H-NS and RpoS dependant manner, and it is postulated that DsrA may activate LEE gene expression through H-NS de-repression at *LEE1* (Bhatt *et al.*, 2011).

1.3.2.9 Non LEE-encoded effector proteins

As previously mentioned, many Gram-negative pathogens utilise a T3SS to inject LEE encoded bacterial effector proteins directly into the cytosol of host cells, to subvert signalling pathways, and promote bacterial colonisation. However, in addition to these LEE encoded effectors, an ever increasing number of non LEE-encoded effector proteins (Nle) have been identified that utilise the

LEE encoded T3SS for translocation (Deng *et al.*, 2004; Gruenheid *et al.*, 2004). In EHEC, Tobe and colleagues (2006) identified 49 putative effectors; including NleA through NleH (including 12 homologues of NleG), EspJ-O, EspR and EspV-Y; many of which are homologues (Dean and Kenny 2009). As most Nles are found on lambda phage and/or transposase like genes, these effectors have likely been acquired by horizontal gene transfer (Iguchi *et al.*, 2009; Tobe *et al.*, 2006). Although these effectors are highly conserved, effector repertoires appear to be flexible between different strains and pathotypes; where EPEC have been shown to harbour a much smaller effector repertoire than EHEC (Iguchi *et al.*, 2009; Tobe *et al.*, 2006).

Many functions have been attributed to these effector proteins. In EHEC and EPEC for example, the effector proteins NleH1 and NleH2 are believed to confer a survival phenotype, where NleHs inhibit apoptosis via various stimuli in epithelial cells, dependant on their binding to the anti-apoptotic Bax inhibitor-1 (Hemrajani *et al.*, 2010; Vossenkamper *et al.*, 2011). Another effector protein attributed with apoptotic activity is the metalloprotease, NleD, which prevents JNK-mediated pro-apoptotic signalling by cleaving and inactivating JNK (Baruch *et al.*, 2010; Vossenkamper *et al.*, 2011). More recently, the effectors

proteins, NleB, NleH, NleC, NleD and NleE, have been demonstrated to harbour anti-inflammatory activity (Nadler *et al.*, 2010; Newton *et al.*, 2010; Royan *et al.*, 2010; Vossenkamper *et al.*, 2010; Vossenkamper *et al.*, 2011; Yen *et al.*, 2010). EspF from EHEC and EPEC prevent phagocytic uptake by macrophages and M cells *in vitro* (Martinez-Argudo *et al.*, 2007; Vossenkamper *et al.*, 2011). Interestingly, this effector is also believed to mediate EPEC invasion into intestinal cells (Weflen *et al.*, 2010; Vossenkamper *et al.*, 2011). Effectors EspH and EspJ have been reported to inhibit phagocytic uptake (Dong *et al.*, 2010; Vossenkamper *et al.*, 2011) and opsonophagocytic uptake respectively (Marches *et al.*, 2008; Vossenkamper *et al.*, 2011).

In *Citrobacter rodentium*, NleA is associated with increased severity of disease (Gruenheid *et al.*, 2004; Schwidder *et al.*, 2011) and NleA has the capacity to compromise protein synthesis after translocation into host cells (Kim *et al.*, 2007; Schwidder *et al.*, 2011). More recently, NleA has been found to play a role in the destruction of tight junctions in EPEC (Thanabalasuriar *et al.*, 2010; Schwidder *et al.*, 2010). As mentioned, the up-regulation of NleA is believed to be mediated via Ler in response to a number of environmental stimuli, such as the SOS response and starvation.

Lastly, *in vivo* analyses have also demonstrated an important role for EspZ in the colonisation of mice by *Citrobacter rodentium* (Deng *et al.*, 2004).

1.3.3 Plasmid encoded factors

All strains of EHEC O157:H7 contain a highly conserved virulence plasmid of approximately 90 kb, termed pO157 (Caprioli *et al.*, 2005). As well as being present in O26:H11 strains, pO157 can also be found in most STEC strains isolated from humans (Nataro and Kaper, 1998). *In vitro* and *in vivo* studies have reported conflicting results on the role of the plasmid in adherence to epithelial cells, where loss of the plasmid has shown both increased adhesion, decreased adhesion, or has had no effect on adhesion (Nataro and Kaper, 1998). However, several reports do correlate pO157 with virulence and it carries a number of putative virulence genes; including four genes encoded on the *Ehx* operon that are necessary for the synthesis and transport of enterohaemolysin (Ehx); a serine protease (EspP); a cluster of 13 genes encoding a type II secretion system; a catalase peroxidase (KatP); a zinc metalloprotease (EscE); and a 9.5 kb putative virulence gene, the product of which has been designated ToxB (Tozzoli *et al.*, 2005).

1.3.3.1 Enterohaemolysin

Plasmid encoded enterohaemolysin was the first pO157 sequence to be determined and hybridisation studies have shown that it is closely related but not identical to α -haemolysin, which is produced by a number of pathogenic *E. coli* strains (Schmidt *et al.*, 1995). Enterohaemolysin belongs to the RTX (repeat in toxin) family of exoproteins, and exerts its toxicity by forming pores in the cytoplasmic membrane of eukaryotic cells and ultimately disrupts the permeability control of the cell (Burland *et al.*, 1998). Belonging to the RTX (repeat in toxin) family of exoproteins, enterohaemolysin exerts its toxicity by forming pores in the cytoplasmic membrane of eukaryotic cells and ultimately disrupts the permeability control of the cell (Burland *et al.*, 1998). Boerlin *et al* (1998) have shown that through restriction fragment length polymorphism (RFLP) of the highly conserved enterohaemolysin (*ehxA*) operon, several active sites for the toxin in an apparently more variable region of the plasmid are also highly conserved. This has led to the suggestion that the toxin is under strong selective constraints and so may play an important role in the survival of EHEC and STEC (Boerlin *et al.*, 1998). Although not proven, enterohaemolysin may aid in the infection process as the iron which is released from lysed erythrocytes (RBC) in the form of haem and haemoglobin may be used by the bacteria to

stimulate growth (Nataro and Kaper, 1998). As the incidence of EHEC enterohaemolysin in STEC strains is high, it has also been implicated as a virulence determinant in HC and HUS. In support of this, serological responses to enterohaemolysin have been frequently observed in humans with O157:H7 associated HUS, but seldom observed amongst controls (Schmidt *et al.*, 1995).

1.3.3.2 EspP

EspP is an extracellular secreted protease which is a member of the serine protease autotransporter of *Enterobacteriaceae* (SPATE) family of exported proteins (Leyton *et al.*, 2003). Although little is known about the source of the *espP* gene, it is widely observed among O157:H7 strains and several homologues have been observed in other pathogenic *E. coli* and enteric bacteria. Examples include the *espC* gene of EPEC, the *pet* gene of EAEC and the *sepA* gene of *Shigella flexneri* (Brunder *et al.*, 1999). Functional analysis of EspP has shown that it is able to cleave pepsin A and human coagulation factor V, and can also induce a cytopathic effect on Vero cells (Leyton *et al.*, 2003). This has led to the suggestion that like enterohaemolysin, EspP may significantly contribute to EHEC pathogenesis in that it probably exacerbates epithelial cell degradation

and haemorrhage following infection, and antibodies to EspP have been found in convalescent-phase sera from children who suffered EHEC infection (Burland *et al.*, 1998; Caprioli *et al.*, 2005).

1.3.3.3 EHEC type II secretion system (T2SS)

The pO157 plasmid contains a cluster of 13 genes (*etpC* through *etpO*) that are closely related to those genes found in the T2SS of Gram-negative bacteria. In *Klebsiella pneumoniae* for example, these genes encode a pullulanase secretion pathway that provides a mechanism for exoprotein secretion (Burland *et al.*, 1998; Caprioli *et al.*, 2005). Although the pO157 plasmid has all the genes necessary for protein secretion, functional export of proteins has yet to be demonstrated (Burland *et al.*, 1998).

1.3.3.4 KatP

KatP is a catalase peroxidase that is encoded on the pO157 plasmid by *katP*. This enzyme may protect EHEC from the reactive oxygen species found in host cells that it will encounter during the infection process (Burland *et al.*, 1998). Interestingly, the occurrence of the *katP* gene in EHEC strains has been shown to

be closely associated with the enterohaemolysin operon found in pO157 (Burland *et al.*, 1998).

1.3.3.5 StcE

StcE is a zinc metalloprotease, the expression of which is regulated by the previously mentioned LEE regulator *ler* (Grys *et al.*, 2005). Encoded by the *stcE* gene on pO157, StcE is able to cleave C1-esterase inhibitor (C1-INH), a serine protease inhibitor that is able to regulate complement and other mechanisms of inflammation (Grys *et al.*, 2005). Interestingly it has been shown that despite being cleaved by StcE, C1-INH can remain functional and as such can continue to inhibit components of the classical complement cascade. Moreover, StcE has been shown to bind and localise functional C1-INH to cell membranes, effectively promoting complement regulation by C1-INH at sites of potential lytic complex formation. Taken together, it has been proposed that StcE potentiates C1-INH activity, thereby reducing inflammation and complement mediated lysis at sites of EHEC infection (Grys *et al.*, 2005). In addition to C1-INH, two other substrates for StcE have been identified, namely glycoprotein 340 (gp340) and mucin 7 (MUC7); both of which have been implicated as

defensive glycosylated proteins in saliva and other tissues (Grys *et al.*, 2005). As StcE exhibits mucinase activity and is also positively regulated by *ler*, the expression of which also regulates many of the effector proteins involved in A/E lesion formation, it has been hypothesised that StcE contributes to the intimate adherence of EHEC to host cells by degrading the protective layers of mucus and glycoproteins (Grys *et al.*, 2005).

1.3.3.6 ToxB

Sequencing of the pO157 plasmid has shown that it contains a large putative virulence gene, termed L7095 in EHEC O157 EDL933 strain and *tox*B in RIMD 0509952 Sakai strain (Tozzoli *et. al.*, 2005). L7095/*tox*B is 9.5 kb in size and its product ToxB shares approximately 20% homology in deduced amino acid sequence with the toxin B of *Clostridium difficile*, and 28% (identical amino acids) and 47% (similar amino acids) homology with the products of the *efa-1/lifA* virulence genes found in non-O157:H7 isolates and EPEC strain 2348/69 (Tozzoli *et al.*, 2005). *efa-1/lifA* has been detected in EHEC strains although these strains contain only a fragment of this gene (Tozzoli *et. al.*, 2005). Tatsuno *et al.*, (2001) have shown that ToxB contributes to the adherence of *E. coli* O157:H7 *in vitro* to

Caco-2 cells through the modulation of proteins secreted by the T3SS. In EPEC strain 2348/69, *lifA* confers the ability to inhibit human peripheral blood lymphocyte proliferation and the mitogen-stimulated synthesis of the cytokines interleukin-2 (IL-2); IL-4; IL-5 and interferon gamma (IFN- γ) (Spears *et al.*, 2006). Although studies have indicated that pO157 encoded *toxB* is consistently present in *E. coli* O157:H7 strains, little is known about its presence in non-O157 and EPEC strains (Tozzoli *et al.*, 2005).

1.3.4 Fimbrial adhesins

E. coli express a variety of different surface factors, including fimbrial and non-fimbrial adhesins, which have been demonstrated to be important for pathogenesis.

1.3.4.1 Type 1 fimbriae

In EHEC O157:H7, at least 16 loci have been identified encoding genes putatively involved in fimbriae or pili biosynthesis (Hayashi *et al.*, 1991; Perna *et al.*, 2001). One of these clusters contain the genes encoding type 1 fimbriae, the most common adhesin produced by *E. coli* which mediates adherence to

mannose-containing glycoproteins found on the surface of eukaryotic cells (Leathart and Gally, 1998). In uropathogenic *E. coli* (UPEC), this adhesin plays a crucial role in the colonisation of the human urinary tract (Holden *et al.*, 2006). Although most EHEC O157:H7 isolates harbour the type 1 fimbrial operon, they are unable to express it due to a conserved 16 bp deletion in the type 1 fimbrial operon promoter region, known as the *fim* switch (Li *et al.*, 1997; Roe *et al.*, 2001). This promoter region is a phase variable invertible DNA element lying upstream of *fimA* encoding the main fimbrial subunit protein, FimA (Abraham *et al.*, 1985). When the *fim* switch is in the off orientation, type 1 fimbriae cannot be produced. This inversion is dependent on the recombinase proteins FimB and FimE, as well as a number of other co-factors (Gally *et al.*, 1996). In EHEC, the 16 bp deletion in the *fimA* promoter region prevents inversion of the *fim* switch to the on orientation, thus type 1 fimbriae cannot be produced (Li *et al.*, 1997; Roe *et al.*, 2001).

1.3.4.2 Long polar fimbriae

Long polar fimbriae (LPF), related in their genetic organisation to type 1 fimbriae, are surface expressed structures that extend from the poles of the bacterial cell (Baumler *et al.*, 1995). First identified in *Salmonella enteric* serovar

Typhimurium, LPF are considered to mediate the attachment of *Salmonella enterica* to Peyer's patch epithelium in mice (Baumler *et al.*, 2006). In the chromosome of EHEC O157:H7 strain EDL933, two non-identical loci encoding LPF (*lpf1* and *lpf2*) have been identified (Perna *et al.*, 2001); and studies indicate that LPF play an important role in the pathogenesis of EHEC and other *E. coli* pathotypes (Jordan *et al.*, 2004; Torres *et al.*, 2002; Newton *et al.*, 2004). In EHEC O157:H7 strain 86-24 for example, isogenic *lpf* mutants have been shown to form fewer A/E lesions *in vivo* and are recovered in significantly lower numbers compared to their parent strain (Jordan *et al.*, 2004). Torres *et al* (2007) have shown that mutation in both the *lpf1* and *lpf2* loci reduces the ability of EHEC O157:H7 to persist in the intestine of lambs. Interestingly, it has also been demonstrated that the same mutant is able to better colonise intestinal mucosal explants *in vitro* compared with the wild-type strain; suggesting that the role of LPF in persistence may not involve initial adherence.

1.3.4.3 Curli

Many *E. coli* strains express a type of fimbriae at their surface known as curli (Barnhart and Chapma, 2006). These thin aggregative fimbriae, or curli fibres, are typically expressed in response to low temperatures (below 30°C), low

oxygen, low nutrients or low osmolarity; as well as during stationary phase growth (Barnhart and Chapma, 2006). For many enterobacteria, curli have been recognised as a major virulence factor due to their role in mediating bacterial cell aggregation, adhesion to surfaces, and biofilm formation (Olsen *et al.*, 1993; Prigent-Combaret *et al.*, 2001); and bind fibronectin, laminin, serum proteins and Congo red dye (Hammar *et al.*, 1995; Ben Nasr *et al.*, 1996; Olsen *et al.*, 1989; Sjobring *et al.*, 1994).

The genes responsible for curli biogenesis are organised into two divergently transcribed operons, *csgBAC* and *csgDEFG*, which are under complex regulation by numerous regulators, including by CsgD itself (Barnhart and Chapma, 2006). CsgD is the major transcriptional activator of the *csgBAC* operon that encodes the curli subunit protein CsgA and the nucleator protein CsgB, which are required for the production of mature curli along with the outer membrane lipoprotein CsgG and two curli assembly factors; CsgE and CsgF (Barnhart and Chapma, 2006). Although encoded by the *csgBAC* operon, CsgC has no known role in curli formation (Barnhart and Chapma, 2006). Lee *et al* (2011) have shown that curli production in EHEC strain EDL933 is an important factor in biofilm formation, along with its ability to produce cellulose. Also, mutations

within the *csgD* promoter have been documented in a few outbreak strains of EHEC O157:H7 and it has been suggested that in these strains, curli expression levels can vary in a temperature-independent phase-variant manner (Uhlich *et al.*, 2001). These examples highlight the importance of curli for biofilm formation in EHEC and the complexity of their regulation, as shown by the markedly different levels of expression that can be observed between different EHEC strains.

1.3.4.4 Type IV pili

EHEC produce a type IV pilus, termed haemorrhagic *E. coli* pilus (HCP) that is composed of a 19 kDa pilin subunit encoded by *hcpA* (Xicohtencatl-Cortes *et al.*, 2007). HCP are important for EHEC adhesion to cultured cells *in vitro*; where inactivation of *hcpA* results in significantly reduced adherence of EHEC to human and bovine intestinal cells (Xicohtencatl-Cortes *et al.*, 2007). More recently, studies have demonstrated that purified HCP has the capacity to significantly increase levels of the pro-inflammatory cytokines IL-8 and TNF-alpha from cultured polarised tissue culture cells; and it is hypothesised that

HCP may play an important role in the pathogenesis of HC following infection by EHEC (Ledesma *et al.*, 2010).

1.3.5 Non-fimbrial adhesins

Iha (iron regulated gene A homologue adhesin) is a 67 kDa outer membrane protein that is capable of conferring an adherence phenotype to non-adherent laboratory *E. coli* strains (Tarr *et al.*, 2000). Sharing a high degree of homology to the IrgA protein of *Vibrio cholerae* (Goldberg *et al.*, 1992), *iha* forms part of a conserved tellurite resistance (Te^R) and adherence-conferring island (TAI) which can be found in distantly related pathogenic *E. coli* (Tarr *et al.*, 2000; Taylor *et al.*, 2002). In EHEC strain EDL933, two copies of *iha* are present in the chromosome on duplicated prophage elements, OI-43 and OI-48 (Perna *et al.*, 2001). In closely related O157:H7 strain Sakai, only one copy of *iha* exists (OI-48 (Hyashi *et al.*, 2001). The TAI itself can be characterised into three functionally distinct gene clusters that encode TE^R, a urease, and the putative adhesins Iha and AIDA-1 (adhesin involved in diffuse adherence). AIDA-I is an auto-transporter membrane protein that confers on EPEC a diffusely adherent phenotype on epithelial cells (Benz and Schmidt, 1989; Benz and Schmidt, 2002). Work by Yin *et al* (2009) has suggested that this cluster of genes may contribute to EHEC

O157:H7 pathogenesis; where deletions in these regions affect bacterial adherence to tissue culture cells and enterocytes in pig ileal loops. The precise roles of Iha and AIDA-1 in EHEC virulence however remain to be clearly demonstrated.

Another protein implicated to have a role in EHEC adherence is the outer membrane protein A (OmpA). Initially identified from a transposon mutant library screened for hyper-adherent phenotypes, an *ompA* mutant was found to adhere poorly to HeLa cells compared with its parent strain. In this same study, antibodies directed against OmpA inhibited adhesion, suggesting OmpA may play an important role during EHEC colonisation. Moreover, recent studies in *Shigella* have provided some evidence that OmpA is an important colonisation factor in this pathogen; where immunisation of mice with recombinant OmpA was enough to protect these animals from experimental challenge with a virulent strain of *Shigella flexneri* 2a (Pore *et al.*, 2011).

1.3.6 Flagella

The genes regulating flagella gene expression and biogenesis are divided among 17 operons which constitute a large and co-coordinately regulated flagellar

regulon. Within this regulon, operons are divided into three temporally-regulated, hierarchical transcriptional classes: early genes (class 1), middle genes (class 2) and late genes (class 3) (Chilcott and hughes, 2000). The early genes encode the master flagella regulator, *flhDC*, which activates transcription of middle genes, including genes encoding the hook-basal body and the transcriptional regulators *flgM* and *fliA* (σ^{28}) (Chilcott and hughes, 2000). *fliA* regulates the transcription of the late genes which include *fliC* and the *mot* operon (Chilcott and hughes, 2000). The flagellar filaments themselves are made up of 11 proto-filaments which polymerise to form monomers of the flagellin protein FliC that are attached to a hook complex which is anchored to a basal body and forming the flagellar motor (MacNab, 2003). The flagellum components are functionally related to the LEE-encoded T3SS and it is hypothesised that flagellin polymerisation occurs via a similar mechanism to the translocation of T3SS effector proteins (Aldridge and Hughes, 2002).

The precise role(s) of flagella (H7) in EHEC pathogenesis are to be properly elucidated although several studies have suggested that flagella are important for EPEC during the early stages of colonisation; where Giron *et al* (2002) demonstrated that a *fliC* mutant is attenuated for adherence *in vitro*. In this

same study however, antibodies directed against EHEC H7 antigen did not inhibit adherence; and purified H7 was shown not to directly adhere to cells (Giron *et al.*, 2002). More recently, studies by Mahajan *et al.* (2009) have provided strong evidence that flagella in EHEC mediate adherence where binding of a *fliC_{H7}* mutant O157 strain to rectal epithelium *in vitro* was significantly reduced; as was binding of the flagellated wild-type strain following incubation with H7-specific antibodies. Moreover, complementation of the *fliC_{H7}* mutant with *fliC_{H7}* was then enough to restore adherence to wild-type levels, suggesting that H7 flagellum acts as an adhesin to bovine intestinal epithelium (Mahajan *et al.*, 2009). Luck *et al.* (2006) have also reported an important role for FliC in EHEC O113:H21; where deletion of *fliC* significantly decreased bacterial invasion on tissue culture cells; a phenotype that was complementable when *fliC_{H21}* was provided back *in trans* on a plasmid. Together, these studies strongly implicate flagella as having an important contribution in the multi-step colonisation process of different *E. coli* strains.

In EHEC, T3S and motility are reciprocally regulated, where T3S is up-regulated and motility is down-regulated in response to different environmental cues; and vice versa. Indeed, EHEC have been shown to shut down flagella gene

expression in micro- colonies whilst up-regulating genes associated with T3S; a switching mechanism that is reported to be mediated via the DNA-bending protein IHF (IHF activates *ler* and represses *flhDC* gene transcription) (Yona-Nadler *et al.*, 2003). Reciprocal regulation of T3S and motility has also been demonstrated for the intrinsic LEE regulators GrlRA; and for a number of QS regulators that are present in the chromosome of *E. coli* (Sperandio *et al.*, 2002; Iyoda *et al.*, 2006).

1.4 Evolution of EHEC

EHEC strains fall into at least four divergent clonal groups. One group includes EHEC strains of serotype O157:H7; a second clonal group includes EHEC strains of serotype O111:H8 and O26:H11; a third group includes EHEC strains of serotype O103:H2 and O45:H2; and a fourth group that includes many different O types that are usually associated with H21 flagellar antigen, such as O113:H21 and OX3:H21 (Caprioli *et al.*, 2005). Evolutionary analysis has shown that *E. coli* O157:H7 is only distantly related to other EHEC clonal groups and is most closely related to an EPEC clone of serotype O55:H7, a non-Stx producing strain of *E. coli* associated with infantile diarrhoea (Whittam *et al.*, 1993). Phylogenetic

studies by Feng *et al* (1998) have inferred that *E. coli* O157:H7 strains evolved in a step wise fashion from an O55:H7 like ancestor by acquisition of bacteriophage derived Stx (Stx1 and Stx2), a large virulence plasmid (pO157), and the transition of somatic antigen O55 to O157. The most common *E. coli* O157:H7 isolates are motile, non-sorbitol fermenting (SOR⁻), and β -glucuronidase negative (GUD⁻), although nonmotile SOR⁺ GUD⁺ O157:H⁻ isolates has been identified in Germany, Hungary, Finland and Swede; and a nonmotile SOR⁻ GUD⁻ O157:H⁻ isolate has been commonly observed in Australia (Steele *et al.*, 2007). Loss of ability to ferment sorbitol and β -glucuronidase activity are believed to be two key landmarks in the proposed stepwise evolution of *E. coli* O157 that has led to the emergence of the most commonly isolated GUD⁻ and SOR⁻ phenotype that accounts for most disease caused by EHEC (Feng *et al.*, 1998; Zhang *et al.*, 2006).

1.5 Comparative genomics: EHEC O157:H7 versus K-12

The complete genome sequence of non-pathogenic laboratory *E. coli* K-12 strain MG1655 was published in 1997 (Blattner *et al.*, 1997). Four years later in 2001, the complete genome sequences of two pathogenic strains of *E. coli* O157:H7

were published; strains EDL933 and Sakai (RIMD0509952) (Hayashi *et al.*, 2001; Perna *et al.*, 2001). On comparison, it was shown that the genome of EDL933 and MG1655 share a large amount of 'backbone' DNA (~ 4.1 million base pairs each) that is arranged similarly in the two strains, with the exception of a 422 kb inversion located at the replication terminus (Blattner *et al.*, 1997; Perna *et al.*, 2001). However, lineage-specific DNA segments were also found to be scattered throughout the genomes in both strains (Perna *et al.*, 2001). In EDL933, these DNA segments (1.34 Mb in total) were numbered and designated 'O-islands' (OIs), and in MG1655, numbered and designated 'K-islands' (KIs) (0.53 Mb in total). In total, 177 OIs and 234 KIs (>50 kb) were identified (Perna *et al.*, 2001). In EDL933, approximately 26% of all the genes in the genome lie within these OIs, and of these OIs, 40% may be assigned a function (Perna *et al.*, 2001). Additionally, ~33% of OIs in the EDL933 genome contain only genes with no known function, 6.2% of which lie within phage-related clusters (Perna *et al.*, 2001). It has been suggested by Perna *et al* (2001) that a large proportion of the DNA found in OIs and KIs was acquired through horizontal gene transfer. Indeed, several of the OIs described by Perna *et al* (2001) include many known and predicted virulence determinants that have been previously well studied, such as the *stx* (*stx1* is located within a cryptic prophage designated CP-933V

and *stx2* is located within a lambdoid phage designated BP-933W) and LEE genes, which are known to be subject to lateral transfer (Whittam and Bumbaugh, 2002).

1.6 Research aims and rationale

The horizontal acquisition of foreign DNA is not only associated with the emergence of pathogenic bacteria and their adaptation to unique niches, but is a major driving force in their diversification and evolution. In the sequenced EHEC O157:H7 strain EDL933, these horizontally acquired elements have been termed O-islands (OIs) and include both fully functional and cryptic prophages. Crucially, comparative genomic analyses of EDL933 with non-pathogenic *E. coli* strain MG1655 have demonstrated that these OIs are specific to the pathogenic *E. coli* strain. Leading on from this, the overall aim of this research was to characterise these prophage-associated regions in EHEC strain TUV93-0, a Shiga toxin-negative derivative strain of EDL933, by phenotypically screening a large collection of OI deletions and comparing these with the parent strain. The rationale was that such phenotypic screening would provide an insight into the function of these islands and the evolution of this specific pathogen, in particular its capacity to colonise the gastrointestinal tract of ruminants and

transmit to humans. Given the focus on screening for effects on T3S, it was envisaged that novel regulators of the LEE PAI may be identified, along with novel OI-derived factors that potentially cross-talk with this essential colonisation system. An aim was to contribute to the understanding of the complex regulatory biology of this pathogen. To achieve this, a variety of molecular genetic and proteomic approaches were adopted that permitted the defined research objectives detailed below:

Objectives

1. Phenotypically screen a large collection EHEC derived OI deletions for important phenotypes such as T3S, motility, growth rate and acid resistance; and then compare these with the parent strain to identify potentially significant OI regions (Chapter 3).
2. Phenotypic and genotypic characterisation of OI regions (identified through the initial phenotypic screening process) using cloning and complementation and allelic exchange mutagenesis in the relevant deletion backgrounds (Chapter 4).

3. Assess the *in vivo* significance of OI targets identified as significant through objectives 1 and 2, by comparing the equivalent deletion strain and wild type parent in a ruminant competition infection model (Chapter 4).
4. Functionally characterise OI encoded proteins of interest to determine their mechanism of action and their interplay with other proteins and regulatory networks in *E. coli* (Chapter 5).

Chapter 2

Materials and methods

2.1 Bacterial strains and plasmids

All of the bacterial strains and plasmids used in this research are detailed in Tables 2.1 and 2.2.

2.2 Bacterial culture conditions and media

Bacterial strains were routinely cultured at 37°C with shaking (200 rpm) in Luria-Bertani (LB) broth or agar (Oxoid) prepared in accordance with the manufacturer's instructions. For the analysis of T3S proteins, minimal essential media with 25 mM HEPES (MEM-HEPES) (Sigma-Aldrich) supplemented with 250 nM Fe(NO₃)₂ and 0.2% glucose was used for culturing strains. For bacterial growth rate assays, M9 medium was used containing 1 × M9 salts supplemented with 2 mM magnesium sulphate (MgSO₄·7H₂O), 0.1 mM calcium chloride (CaCl₂·H₂O) and 0.2% glucose. CFA agar (LabM), prepared according to the manufacturer's instructions, and CFA broth (0.15% yeast extract (Becton Dickinson), 1% casamino acids (Oxoid), 0.01% Congo red (Sigma-Aldrich), 400 µM magnesium sulphate and 40 µM manganese chloride (MnCl₂·4H₂O)) was used for selecting and culturing colonies respectively in Congo-red binding assays. When required, Isopropyl β-D-1-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside (X-gal) and antibiotics were

added to media at the follow final concentrations: 1 mM IPTG; 20 ng/ml⁻¹ X-gal; 50 µg/ml⁻¹ ampicillin; 50 µg/ml⁻¹ chloramphenicol; 100 µg/ml⁻¹ kanamycin and 15 µg/ml⁻¹ nalidixic acid.

2.3 DNA analysis and genetic manipulation methods

Unless otherwise stated, all methods were performed as for Sambrook *et al* (1989).

2.3.1 Genomic DNA extraction from *E. coli* strains

Genomic DNA was extracted from *E. coli* isolates using a ChargeSwitch genomic DNA (gDNA) Mini Bacterial Kit (Invitrogen) in accordance with the manufacturer's instructions. Aliquots of purified DNA were stored at -20 °C and genomic DNA was diluted 1:100 in Milli-Q water (MQ) (Sigma-Aldrich) before being used as DNA template in Polymerase Chain Reaction (PCR).

2.3.2 Preparation of crude DNA extracts from *E. coli* strains

Crude DNA preparations were used as templates for diagnostic and screening PCRs mainly in cloning experiments. A single bacterial colony from overnight growth on LB agar was suspended in 100 µl MQ water and heated in a heat

block to 100°C for 10 min. Aliquots were stored at -20°C and 1 µl of lysate was used as DNA template in all PCR reactions.

2.3.3 Plasmid DNA extraction from *E. coli* strains

Plasmids were extracted from bacteria cultured overnight in 5 ml LB broth supplemented with appropriate antibiotic using the GeneJET Plasmid Miniprep Kit (Fermentas) in accordance with the manufacturer's recommendations.

2.3.4 DNA amplification by PCR

All of the PCR primers used in this research are detailed in Table 2.3. PCR was carried out in sterile thin-wall tubes in 50 µl reaction volumes, containing 1 µl template DNA (1:100 dilution for gDNA), 200 µM dNTP mixture (Roche), 0.5 µM primer, 1 × ThermoPol *Taq* PCR reaction buffer (20 mM Tris-HCl, 10 mM (NH₄)₂SO₄, 10 mM KCl, 2 mM MgSO₄, 0.1% Triton X-100 pH 8.8 at 25 °C), 1.25 U *Taq* DNA polymerase (NEB) and MQ water to the final volume. For high-throughput PCRs (e.g. screening bacterial colonies for positive clones), Quick-Load *Taq* 2× Master Mix (NEB) was used according to the manufacture's guidelines and contained 1.0 µl template DNA and 0.5 µM of each primer.

For cloning and sequencing purposes, high fidelity Expand Long Template PCR (Roche) was used to amplify large DNA fragments (>4 Kb) and Phusion high-fidelity DNA polymerase (Finnzyme) was used to amplify DNA fragments <4 Kb; both according to the manufacturer's guidelines and contained 1 μ l template DNA, 200 μ M dNTP, 0.5 μ M each primer, 1 \times kit specific reaction buffer, 1.5 U *Taq*/*Tgo* polymerase mix and 0.5 U Phusion DNA polymerase respectively. High Fidelity Expand Long Template PCR contains an enzyme mix of *Taq* DNA polymerase and *Tgo* DNA polymerase with inherent proofreading activity, ensuring a lower error rate than PCR using *Taq* DNA polymerase alone. Phusion High-Fidelity DNA Polymerase contains a *Pyrococcus*-like enzyme and has an error rate 50-fold lower than *Taq* DNA polymerase. All PCRs were performed in a Thermo-Hybaid PCR Express cyclor using cycling conditions in line with each supplier's specific recommendations. Amplification products were subsequently resolved by agarose gel electrophoresis as described in section 2.11.2

2.3.5 DNA electrophoresis

DNA fragments were resolved by electrophoresis in gels containing between 0.8-1.2 % agarose in 1 \times Tris-borate/EDTA (TBE) buffer (89 mM Tris-borate and 2

mM ethylenediamine tetraacetic acid (EDTA) (Sigma-Aldrich), pH 8.3, containing ethidium bromide or SafeView (NBS Biologicals) for DNA staining. Electrophoresis was performed in 1 × TBE buffer at constant voltage (100 V) using Bio-Rad systems. DNA samples were mixed with DNA loading buffer (Invitrogen) prior to loading and fragment sizes were determined by comparison to molecular markers (NEB; 1 Kb or 100 bp). PCR amplicons were visualised under UV illumination in a Flowgen MultiImage light cabinet (Shenstone, England) and images were captured and analysed using ChemiImager 4000i v.4.04 software.

2.3.6 DNA fragment purification

2.3.6.1 Column and gel purification

DNA fragments from PCR and restriction digests were purified using the GeneJET PCR Purification Kit (Fermentas) and DNA fragments from agarose gels were purified using the GeneJET Gel Extraction Kit (Fermentas); according to the manufacturer's instructions.

2.3.7 Transformation of competent *E. coli*

E. coli strains were made competent for transformation with plasmid DNA using the protocols described in Sambrook *et al.*, 1989.

2.3.8 DNA digestion

Enzymatic digestion of DNA was carried out in a 50 µl reaction volume in a 1.5 ml eppendorf tube (Greiner) and typically contained 30 µl genomic or plasmid DNA, 2 µl restriction endonuclease (NEB), 5 µl 10 × appropriate NEB buffer, 5 µl 10 × bovine serum albumin (BSA) (NEB) and MQ water to the final volume. Incubation was carried out for at least 2 hours in a water bath set at the appropriate temperature for each restriction enzyme.

2.3.9 Klenow fill-in

When necessary, digested DNA products were blunt-ended using Klenow (Large fragment of DNA polymerase I) (NEB) in a 60 µl reaction volume in a 1.5 ml eppendorf tube (Greiner) and typically contained 30 µl plasmid DNA, 2 µl Klenow fragment (NEB), 6 µl 10 × Klenow reaction buffer, 6 µl 10 × bovine serum albumin (BSA) (NEB) and MQ water to the final volume. Incubations were carried out at RT for 30 min and then the reactions were stopped by

column clean up as described in section 2.3.6.1. Subsequent ligation reactions were carried out exactly as described in section 2.3.10.

2.3.10 DNA ligation

In the construction of recombinant DNA plasmids, vector and insert DNA was digested with the appropriate restriction endonucleases and purified as described in section 2.3.6.1. Linearised vector was treated with shrimp alkaline phosphatase (SAP) (NEB) to dephosphorylate 5' restricted termini to help prevent vector re-ligation. Ligation reactions were carried out in a thin-walled tube and typically contained 6 µl insert, 2 µl vector, 1 µl 10 × ligase reaction buffer (50 mM Tris-HCl, 10 mM MgCl₂, 1 mM ATP, 10 mM dithiothreitol (DTT), 25 µg/ml BSA, pH 7.5 at 25°C) (NEB) and 1 µl T4 DNA ligase (NEB) to a final volume of 10 µl. Ligation reactions were mixed thoroughly and incubated in a Thermo-Hybaidd thermocycler machine at 16°C overnight.

2.3.11 Preparation and transformation of chemically-competent *E. coli* cells

E. coli strain DH5α (Invitrogen) was cultured in 5 ml LB broth shaking at 200 rpm at 37°C overnight. The following day the culture was diluted 1:100 in LB broth and cultured under the same conditions until bacterial growth reached an

optical density at 600 nm (OD_{600}) of between 0.5-0.7. Cultures were incubated on ice for 20 min prior to centrifugation at $4000 \times g$ for 8 min at 4°C. Cell pellets were gently suspended in 0.4 volumes (of the original culture volume) of ice-cold transformation buffer I (TFBI: 30 mM potassium acetate ($KC_2H_3O_2$), 10 mM $CaCl_2 \cdot 2H_2O$, 100 mM KCl, 15% (v/v) glycerol and 45 mM $MnCl_2$) and incubated on ice for 15 min. Cells were harvested as above and gently suspended in 0.04 volumes of ice-cold transformation buffer II (TFBII: 75 mM $CaCl_2$, 10 mM KCl, 15% (v/v) glycerol and 10 mM Na-MOPS pH 7.0). Aliquots (200 μ l) of cells were stored at -70°C.

Chemically-competent *E. coli* DH5 α cells were defrosted on ice and mixed with 10 μ l of ligation reaction (sections 2.3.10 and 2.3.11). The cell-ligation mixture was transferred into a glass tube and incubated on ice for 30 min before heat shocking in a water bath at 42°C for 45 sec. Reactions were incubated on ice for a further 2 min and then 800 μ l of SOC media (Invitrogen) was added. Transformants containing pIB307-based temperature-sensitive vectors were recovered at 28°C for 3 h with shaking at 100 rpm. All other plasmid transformants were recovered by incubation at 37°C for 90 min with shaking (200 rpm). 200 μ l of recovered culture was plated onto LB agar supplemented

with the appropriate antibiotics and incubated overnight at the appropriate temperature.

2.3.12 Preparation and transformation of electro-competent cells

All plasmid transformations in wild-type *E. coli* and strains and their derivatives were performed by electroporation. Preparation of cultures to make electro-competent cells was in a similar manner as for chemically-competent cells (section 2.3.11). Culture pellets were washed twice by suspending in 0.5 volumes of ice-cold 10% glycerol (v/v) (Fisher Scientific) and then centrifuging at $4000 \times g$ for 8 min at 4°C. The final cell pellet was resuspended in 0.005 volumes of ice-cold 10% glycerol (v/v) and stored in aliquots (50 µl) at -70°C.

Electro-competent cells were defrosted on ice and mixed with 1 µl purified plasmid DNA. The cell-plasmid mixture was transferred into a pre-chilled 1.5 ml electroporation cuvette (Flowgen, UK) and electroporated at 2.5 KV in a (machine name). SOC medium (1 ml) was added immediately and the suspension was transferred into a 1.5 ml eppendorf tube and transformants were recovered as described for the transformation of chemically competent cells (section 2.3.11). 80 µl of recovered culture was plated onto LB agar

supplemented with appropriate antibiotic and incubated overnight at the appropriate temperature.

2.3.13 TOPO TA cloning

The TOPO TA cloning kit (Invitrogen) provides an efficient strategy for the direct insertion of *Taq* DNA polymerase amplified PCR products into commercial cloning vector pCR4-TOPO. *Taq* amplified PCR products contain poly-A 3' overhangs that are created by the addition of single deoxyadenosine (A) to the 3' end of PCR products during amplification as a result of *Taq*'s non-template dependant terminal transferase activity. pCR4-TOPO vectors contain poly-T 3' overhangs that increase the efficiency of vector ligation with *Taq* DNA polymerase amplified PCR products carrying complementary poly-A 3' overhangs. Ligation reactions were carried out according to the manufacturer's instructions (Invitrogen TOPO TA user manual version U) and chemically transformed into One shot TOP10 chemically-competent *E. coli* cells (Invitrogen); also according to the manufacturers instructions.

2.3.14 Construction of OI deletions in EHEC strain TUV93-0

All pre-study OI deletion mutants were constructed in the laboratory of Professor John Leong, University of Massachusetts, using the method described by Campellone *et al* (2004).

2.3.15 Construction of defined mutations in EHEC and UPEC

Deletions were constructed for genes *ecs1581* (EHEC), *qseC* (EHEC) and *c1493* (UPEC) using allelic exchange as described by Blomfield *et al* (1991) and Emmerson *et al* (2006). Briefly, chromosomal regions (~600-1000bp) flanking the gene of interest were PCR amplified with a high fidelity Phusion polymerase (Fiinzyne) (Section 2.3.4) and cloned into a chloramphenicol resistant temperature sensitive exchange plasmid (pIB307). A kanamycin cassette was then cloned in between the homologous flanking regions on the exchange plasmid and subjected to several rounds of temperature and antibiotic selection. Integration of the exchange plasmid was achieved at 42°C in the presence of chloramphenicol, whilst plasmid excision was achieved at 28°C in the presence of kanamycin. All mutants were verified by PCR and successful complementation was achieved by providing the related genes *in trans* on low copy number plasmid pWSK29.

2.3.16 Construction of plasmids for complementation

All complementation plasmids were constructed using PCR products amplified with high fidelity Phusion polymerase cloned into pWSK29. *E. coli* strain DH5 α was used as the intermediate host strain for cloning constructs and all constructs were verified by sequencing (GenePool).

2.3.16.1 Construction of engineered ECs1581 and ECED1_1787 protein variants

Site directed mutagenesis was carried out using an overlap PCR approach. Briefly, the 5' and 3' regions of each gene were PCR amplified with engineered mutations in the relevant primers. The two PCR products were then used as a template for the overlap PCR using the outside primers. The PCR products were cloned into vector pWSK29, verified by sequencing, and then tested for their effects on motility and T3S in an *ecs1581* deletion strain background.

2.3.17 Measurement of LEE1 promoter activity

In order to assess LEE1 gene expression in both wild type and mutant *E. coli* O157:H7 backgrounds, a 428 bp PCR generated DNA promoter fragment (-444 to -16 bp upstream from the *ler* ATG start codon) was amplified from strain TUV93-0 using high fidelity Phusion polymerase, and cloned into the promoter-

less green fluorescence protein (*gfp*) plasmid pKC26 to create transcriptional fusion plasmid p*LEE1::gfp*. Test strains harbouring this reporter were cultured in DMEM supplemented with chloramphenicol and the GFP produced by each bacterial population was measured every 60 min by transferring 200 µl aliquots of culture into triplicate wells in a black 96-well plate (FluoroNunc) and reading the plate in a fluorimeter (FLUOstar Optima) using 485 nm absorbance and 520 nm emissions at a gain of 1250. Promoter-less plasmid pKC26 in each strain background acted as a control for auto-fluorescence. When *LEE1* promoter activity was measured in response to induced expression of *ecs1581* from low copy number plasmid pWSK29 (pECs1581), strains were cultured in DMEM supplemented with chloramphenicol, ampicillin and IPTG and both pKC26 and pWControl in each strain acted as a control for auto fluorescence. All were tested in triplicate and on at least three separate occasions.

2.3.18 RT-qPCR

Total RNA was purified from bacteria using an RNeasy kit (Qiagen) and random primers were used to reverse transcribe the RNA using Affinityscript (Stratagene); both in accordance with the manufacturer's guidelines. qPCR was carried out with a PowerSybr mastermix (Applied Biosystems) and amplified in

a MxPro 3000 qPCR machine (Stratagene). Transcript abundance was normalized to 16SrRNA and relative transcription calculated using MxPro software (Stratagene).

2.4 Bioinformatic analyses

Variants of ECs1581 were identified using blastp (www.ncbi.nlm.nih.gov) and sequences were obtained from Genbank (www.ncbi.nlm.nih.gov/genbank). Putative hypothetical domains in ECs1581 were identified using MyHits (www.myhits-isb-sib.ch) and KEGG (<http://www.genome.jp/kegg>). OI-51 prophage alignments were adapted from Magnifying Genomes (MaGe), a relational based prokaryotic genome database (www.genoscope.cns.fr). Sequence alignments were performed using ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2>), EMBOSS Align (<http://www.ebi.ac.uk/Tools/emboss/align>) and LALIGN (http://www.ch.embnet.org/software/LALIGN_form.html). CLC Sequence Viewer (<http://www.clcbio.com>) was used to construct and visualize alignments among the protein coding regions of ECs1581; the ECs1581 region was identified in un-annotated draft *E. coli* sequences through BLASTN searches.

2.5 Sequencing

All DNA sequencing was carried out by GenePool, School of Biological Sciences, University of Edinburgh, and was performed using an ABI Prism BigDye terminator cycle sequencing kit version 3.1 (Applied Biosystems). Reactions were analysed on an ABI 3730 DNA sequencer.

2.6 Bacterial growth analysis

A single bacterial colony for each test strain was cultured overnight in 5 ml M9 media at 37°C with shaking (200 rpm). Cultures were diluted in the same media to a starting OD₆₀₀ of 0.05, incubated as before and OD₆₀₀ measurements taken every 60 min (Cecil spectrophotometer). The mean generation time (g) for each strain was then calculated (using the equation $\log_{10} N_t = \log_{10} N_0 + g \log_{10} 2$) and plotted as a bar chart.

2.7 Acid survival

Ability to survive in acid was assessed using the method described by Small *et al* (1994). In brief, strains were grown overnight in LB broth at 37°C with shaking (200 rpm) before being diluted 10⁻³ in LB broth adjusted to pH 2 with HCL.

Strains were incubated as before for 2 h and serial dilutions were made in sterile phosphate buffered saline (PBS) and plated in triplicate onto LB agar (100µl/plate) and incubated overnight at 37°C. Colony counts (cfu/ml) were compared with those from plated dilutions of the original culture to determine the percentage survival of each strain following acid challenge.

2.8 Congo red binding assay

Strains were initially streaked onto CFA agar plates (LabM) before individual colonies were taken and inoculated into CFA broth and grown overnight at 37°C with shaking. Cultures were diluted in PBS to an OD₆₀₀ of 1.0, centrifuged at $4,000 \times g$ for 5 min and re-suspended in 1 ml of sterile PBS. Washing was repeated a further two times to remove all traces of culture broth before finally re-suspending the culture pellets in 2.5 ml of PBS containing a final concentration of 0.01% Congo red. Samples were incubated at 37°C for 30 min before a 1 ml aliquot was taken and centrifuged at $10,000 \times g$ for 10 min. Supernatants were discarded and pellets were air dried and photographed using a Lumina digital camera. All strains were tested in triplicate and on three separate occasions.

2.9 Motility assays

Bacterial motility was assessed using a soft agar method. Motility plates were prepared by adding 0.3% agar (Sigma-Aldrich) to 1% Bacto-Tryptone (Becton Dickinson) broth containing 0.5% NaCl (Sigma-Aldrich); or by adding autoclaved agar to pre-warmed DMEM to give a final agar concentration of 0.3%. For plasmid based complementation experiments, IPTG and ampicillin was added to plates when appropriate and all plates were poured the night before use and allowed to air dry on the bench. Sterile inoculation needles were used to touch the tops of isolated colonies and stabbed vertically into the agar making sure not to touch the bottom of the plates to avoid twitching activity. Plates were incubated either at 28°C or 37°C for 16 h and motility was assessed by observing the circular halo that forms as a result of the bacteria swimming through the agar. All strains were tested in quadruplicate and each experiment was carried out in triplicate.

2.10 *In vivo* colonisation model

Crossbred lambs (six weeks old) were housed in bio-secure containment level 2 accommodation and supplied with food and water *ad libitum*. All lambs were confirmed to be free of EHEC O157 by enrichment and O157 immuno-magnetic

separation prior to commencing the studies. Lambs were dosed orally with 1×10^{10} CFU of WT *E. coli* O157:H7 and mutant bacteria (5×10^9 CFU of each). Inocula (10 ml resuspended in 10 ml of PBS pH 7.4) were delivered using a worming gun (Novartis Animal Health) ensuring that the whole inoculum was delivered directly to the pharynx. Rectal faecal samples from each lamb were collected daily for direct plating onto sorbitol-MacConkey (Oxoid) plates supplemented with appropriate antibiotics. Enrichment was carried out on samples in buffered peptone water for 6 h at 37°C and then plated onto sorbitol-MacConkey plates supplemented with the appropriate antibiotic. Representative colonies were confirmed to be *E. coli* O157 by latex agglutination (Oxoid). Animal experiments were performed in line with the Animals Scientific Procedures Act (1986) and were approved by the local ethical review committee.

2.11 Methods for protein analysis

2.11.1 Preparation and analysis of T3S culture supernatant proteins

Strains were cultured overnight in either 5 ml LB broth or MEM-HEPES with added glucose and iron at 37°C with shaking (200 rpm) before being diluted to an optical density at OD₆₀₀ of 0.05 in the same media. Cultures were incubated

as previously described and grown to a final OD₆₀₀ of 1. Cultures were centrifuged at 4,000 × g for 30 min at 4°C and supernatants were filtered through 0.45µm low protein-binding filters (Millipore). A 10% (v/v) final concentration of trichloroacetic acid (TCA; Sigma-Aldrich) was used to precipitate the proteins and bovine serum albumin (BSA; NEB) (4µg/ml) was added to act as a co-precipitant and to rule out the precipitation procedure as a source of variation in protein levels between culture supernatants. Supernatants were incubated overnight at 4°C and centrifuged at 4,000 g for 30 min at 4°C. Protein pellets were air-dried and re-suspended in an appropriate volume of re-suspension buffer (1.5 M Tris-HCL) to standardise samples and take into account the slight variation in OD₆₀₀ at which cultures were harvested. Culture supernatant proteins were subsequently analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 12% polyacrylamide gel with Colloidal blue staining (Severn Biotech). In order to verify the presence of type three associated proteins in culture supernatants, antibody detection of EspD (T3 secreted translocon protein) was carried out as described in section 2.11.4 and was used semi-quantitatively gauge T3 levels.

2.11.2 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

Secreted or whole cell protein samples were separated by SDS-PAGE using either BIO-RAD Mini-protean II gel apparatus (Bio-Rad) or a Hoefer SE600 Midi gel electrophoresis unit. Protein samples were mixed 1:1 with 2 × SDS loading buffer (Sigma-Aldrich) and incubated at 100°C for 5 min to denature the proteins. Protein samples were loaded in a 5% stacking gel (5% acrylamide-bisacrylamide, 0.125 M Tris-HCl, pH 6.8, 0.1% SDS, 0.1% ammonium persulfate (APS) and 0.001% N,N,N',N'-Tetramethylethylenediamine (TEMED)) and resolved in a 12% resolving gel (12% acrylamide-bisacrylamide, 0.375 M Tris-HCl, pH 8.8, 0.1% SDS, 0.1% APS, 0.001% TEMED) by electrophoresis. Electrophoresis was carried out in Tris-glycine running buffer (25 mM Tris-HCl, pH 8.3, 250 mM glycine and 0.1% SDS) at a constant voltage of either 100 V or 150 V for mini and midi gels respectively. Protein sizes were determined using a broad range pre-stained protein marker (NEB colorplus; 7-174 kDa).

2.11.3 Colloidal blue staining of SDS-PAGE gels

Proteins separated by polyacrylamide gel electrophoresis were stained overnight with Colloidal blue (Severn Biotech) according to manufacturer's instructions. Gels were de-stained in distilled water and images were captured

using a Flowgen MultiImage light cabinet and ChemiImager 4000i v.4.04 software.

2.11.4 Protein detection

SDS-PAGE separated proteins (section 2.11.2) were transferred onto Hybond ECL nitrocellulose membrane (Amersham Biosciences) using a Trans-Blot electrophoretic transfer cell (Bio-Rad). Briefly, the gel and nitrocellulose membrane were placed between two pieces of filter paper (3MM Whatman paper), which were in turn placed between two sponge pads, pre-soaked in transfer buffer (20 mM Tris, 154 mM glycine and 20% (v/v) methanol) and held together in a plastic cassette. Transfer was carried out in cold transfer buffer at a constant voltage of 60 V for 90 min. Nitrocellulose membranes were blocked with 8% (w/v) dried milk powder (Marvel) in PBS at 4°C overnight and incubated with the relevant antibody diluted in wash buffer (1% dried milk (Marvel) and 0.05% (v/v) polyoxyethylenesorbitan monolaurate (Tween 20, Sigma-Aldrich) in PBS): mouse monoclonal anti-EspD and rabbit polyclonal anti-mouse IgG HRP conjugated antibodies (Dako) were both diluted 1:4000 and mouse monoclonal anti-His (Qiagen) antibody was diluted 1:500. All antibody-membrane incubations were carried out for 1h at RT on a platform

shaker and were washed for 3×10 min in wash buffer before and after each antibody step. Detection of His-tagged proteins was carried out according to Protocol 7 in the QIAexpress detection and assay handbook and proteins were sized using a His-tagged protein ladder (6×HPL; Qiagen) which also acted as a control for the successful transfer of His-tagged proteins onto the nitrocellulose membranes. For ECL detection, membranes were incubated in 2.5 ml of ECL Solution 1 (100 mM Tris-HCl, pH 8.5, 2.5 mM Luminol ($C_8H_7N_3O_2$) and 0.4 mM *p*-Coumaric acid ($C_9H_8O_3$)) mixed with 2.5 ml of ECL Solution 2 (100 mM Tris-HCl, pH 8.5 and 0.02% hydrogen peroxide (H_2O_2)) for 5 minutes at room temperature on the bench. Chemiluminescence was detected on Hyperfilm ECL chemiluminescence film (Amersham Bioscience) developed in a Protec automatic film processor (Optimax). Images were taken using a Flowgen MultiImage light cabinet and ChemiImager 4000i v.4.04 software.

2.11.5 Native protein purification

E. coli BL21 (DE3) (Invitrogen) cells harbouring plasmid pET28a-ECs1581, pET28a-ECED1_1787 or pET28a-ECED1_1787C20R were grown in LB broth in the presence of kanamycin at 37°C with shaking to an OD_{600} of 0.5. Cultures were induced with IPTG and grown for a further 4 h under the same conditions

before harvesting by centrifugation at $4000 \times g$ for 30 min. Cell pellets were suspended in lysis buffer (50 mM NaH_2PO_4 , 300 mM NaCl, 10 mM imidazole and 1 mg/ml lysozyme, pH 8.0) and lysed by sonication (6 \times 10 second pulses with a 10 second cooling period on ice in between each pulse). Lysates were centrifuged at $10,000 \times g$ at 4°C for 30 min, incubated with Ni^{2+} NT resin (Qiagen) for 1 h at RT on a platform shaker and loaded onto a purification column (BioRad). The column was washed twice with wash buffer (50 mM NaH_2PO_4 , 300 mM NaCl, 20 mM imidazole, pH 8.0) and the protein was eluted from the column in elution buffer (50 mM NaH_2PO_4 , 300 mM NaCl, 250 mM imidazole, pH 8.0). Crude preparations were transferred into dialysis tubing (8,000 kDa molecular weight cut-off, Medicell international Ltd) and dialysed in 2 L of sterile PBS which was replenished every 24 hours for a total of three changes. Total protein concentration was determined by BCA according to the manufacturer's instructions (Novagen) and purity was verified by SDS-PAGE and antibody detection of recombinant His-tagged proteins. Purified proteins were subsequently used in electrophoretic mobility shift and *in vitro* transcription assays along with PCR generated DNA promoter fragments as described in sections 2.12 and 2.13.

2.12 Electrophoretic mobility shift assay (EMSA)

To determine if ECs1581 activation of LEE1 is direct or indirect, a 291 bp LEE1 promoter (-288 to +3, with respect to the *ler* ATG start codon) and a control *gapA* fragment (also 291 bp) were PCR amplified from EHEC strain TUV93-0 and subjected to EMSA with either purified ECs1581, ECED1_1787 or ECED1_1787 C20R proteins. EMSA was carried out as described previously (Tree *et al.*, 2011). Briefly, the purified proteins were co-incubated with concentrations of ddUTP-11-DIG (Roche) end-labeled DNA in binding buffer (10mM Tris (pH 8.0), 50mM KCl, 5 mM EDTA, 200µg/ml BSA, 100ng/ml poly d(I-C), 50mM glutamate and 5% glycerol) for 30 min at 25°C and loaded onto 5% non-denaturing polyacrylamide gels in 0.5X TBE. 50 ng of unlabelled DNA was used for competition experiments and added to the binding reaction. The DNA and DNA-protein complexes were electro-transferred onto nylon membranes and developed with AP conjugated anti-DIG antibody (Roche) as per manufacturer's instructions.

2.13 *In vitro* transcription assay (IVT)

In vitro transcription experiments were performed as described previously (Schneiders *et al.*, 2006). Briefly ×5 IVT Buffer, 2nM PCR product of the test and

control promoters (*LEE1* and *gapA*), RNA polymerase (Epicenter, Madison, WI), RNAsin and test proteins, ECs1581 and/or H-NS (gift of Prof. C. Dorman; Tupper *et al.*, 1994), were incubated for 15 min at 37°C prior to the addition of the transcription mix containing x5 IVT buffer, heparin (1.2 µg/ml), NTPS, and $\alpha^{32}\text{P}$ -UTP (Perkin Elmer, UK). The reaction was stopped 5 min later followed by the addition of Gel loading buffer II (Ambion) and denaturation at 95°C for 5 min. The resulting products were electrophoresed on a 7% polyacrylamide/8M urea gel. Images were acquired following exposure of the dried gels to a phosphorscreen.

2.14 Data analyses

Data was analysed and graphs were drawn using Microsoft® office Excel 2003. Significant differences were assessed by the Student's t-test using Minitab software. All error bars that are shown are a result of the standard deviation of the mean. Statistical advice was provided by Dr Darren Shaw, Veterinary Clinical Sciences, R(D)SV, University of Edinburgh, Easter Bush Veterinary, Roslin, Midlothian who plotted and analysed the ruminant colonisation data shown in Fig. 4.1.

Table 2.1 List of strains used in this research

Strain	Description	Source/Reference
TUV93-0	Shiga toxin negative derivative of EHEC O157:H7 strain EDL933	Campellone <i>et al.</i> , 2004
TUV93-0 Nal ^R	Naladixic acid resistant strain of TUV93-0	This study
TUV93-0 Δ OI-1Cam ^R	Lambda red generated OI-1 chloramphenicol cassette replacement mutant in strain TUV93-0	John Leong, University of Massachusetts, USA Campellone <i>et al.</i> , 2004
TUV93-0 Δ OI-7 Cam ^R	Lambda red generated OI-7 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-8 Cam ^R	Lambda red generated OI-8 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-10-13 Cam ^R	Lambda red generated OI-10-13 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-14-15 Cam ^R	Lambda red generated OI-14-15 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-28-29 Cam ^R	Lambda red generated OI-28-29 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-30 Cam ^R	Lambda red generated OI-30 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-35 Cam ^R	Lambda red generated OI-35 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-36 Cam ^R	Lambda red generated OI-36 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-44 Cam ^R	Lambda red generated OI-44 mutant in strain TUV93-0	John Leong

Table 2.1 strains *cont*

TUV93-0 Δ OI-47 Cam ^R	Lambda red generated OI-47 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-50 Cam ^R	Lambda red generated OI-50 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-51 Cam ^R	Lambda red generated OI-51 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-52 Cam ^R	Lambda red generated OI-52 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-61 Cam ^R	Lambda red generated OI-61 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-62-63 Cam ^R	Lambda red generated OI-62-63 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-66 Cam ^R	Lambda red generated OI-66 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-67 Cam ^R	Lambda red generated OI-67 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-70 Cam ^R	Lambda red generated OI-70 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-76 Cam ^R	Lambda red generated OI-76 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-79 Cam ^R	Lambda red generated OI-79 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-80 Cam ^R	Lambda red generated OI-80 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-102 Cam ^R	Lambda red generated OI-102 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-112-113 Cam ^R	Lambda red generated OI-112-113 mutant in strain TUV93-0	John Leong
TUV93-0 Δ OI-115 Cam ^R	Lambda red generated OI-115 mutant in strain TUV93-0	John Leong

Table 2.1 strains *cont*

TUV93-0 ΔOI-122 Cam ^R	Lambda red generated OI-122 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-133 Cam ^R	Lambda red generated OI-133 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-138 Cam ^R	Lambda red generated OI-138 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-139 Cam ^R	Lambda red generated OI-139 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-141 Cam ^R	Lambda red generated OI-141 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-144 Cam ^R	Lambda red generated OI-144 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-148A Cam ^R	Lambda red generated OI-148A mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-148C Cam ^R	Lambda red generated OI-148C mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-149-150 Cam ^R	Lambda red generated OI-149-150 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-153 Cam ^R	Lambda red generated OI-153 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-154 Cam ^R	Lambda red generated OI-154 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-156 Cam ^R	Lambda red generated OI-156 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-172 Cam ^R	Lambda red generated OI-172 mutant in strain TUV93-0	John Leong
TUV93-0 ΔOI-173-175 Cam ^R	Lambda red generated OI-173-175 mutant in strain TUV93-0	John Leong

Table 2.1 strains *cont*

TUV93-0 $\Delta ecs1581$ (Kan ^R)	Allelic exchange strain with a kanamycin resistance cassette inserted into the open reading frame of <i>ecs1581</i> in strain TUV93-0	This study
TUV93-0 $\Delta qseC$ (Kan ^R)	Allelic exchange strain with a kanamycin resistance cassette inserted into the open reading frame of <i>qseC</i> in strain TUV93-0	This study
CFT073	UPEC pyelonephritis strain CFT073 (serotype O6:K2:H1)	Mobley <i>et al.</i> , 1990
CFT073 $\Delta c1493$ (Kan ^R)	Allelic exchange strain with a kanamycin resistance cassette inserted into the open reading frame of <i>c1493</i> in strain CFT073	This study
ZAP198	Wild type EHEC O157:H7 strain	Low <i>et al.</i> , 2006
ZAP198 $\Delta grlA::Tn5$ (Kan ^R)	<i>GrlA</i> transposon mutant in EHEC strain ZAP198	This study
ZAP1327 (Kan ^R)	Constitutive LEE1 expression strain with a kanamycin resistance cassette inserted into the LEE1 promoter region in strain ZAP198	This study
ZAP1004	Unmarked <i>ler</i> deletion mutant in strain ZAP198	Low <i>et al.</i> , 2006

Table 2.1 strains *cont*

ED1a	Commensal <i>E. coli</i> strain (serotype O81) isolated from the faeces of a healthy individual	Clermont <i>et al.</i> , 2008
DH5 α	Host strain used for cloning with pWSK29	Invitrogen
BL21 (DE3)	Bacteriophage T7 expression <i>E. coli</i> strain used with pET28a (+) to make recombinant His-tagged proteins	Invitrogen

Table 2.2 Plasmids used in this research

Plasmid purpose/name	Description	Source/Reference
Cloning plasmids		
pET28a (+) (Kan ^R)	His-tagged expression vector	Novagen
pKC26 (Cam ^R)	Promoter-less GFP+ reporter fusion construct	Holden <i>et al.</i> , 2007
pWSK29/pWControl (Amp ^R)	IPTG inducible low copy number cloning vector	Wang and Kushner, 1991
pIB073 (Cam ^R)	Chloramphenicol resistant temperature sensitive plasmid used for allelic exchange	Blomfield <i>et al.</i> , 2001
pDG28 (Cam ^R Kan ^R)	Host vector harbouring the kanamycin resistance cassette used for all pIB073 based allelic exchanges	Gally <i>et al.</i> , 1994
Plasmid complementation		
pZ1835-Z1843	<i>z1835-z1843</i> cloned into low copy number plasmid pWSK29	This study
pZ1841	<i>z1841</i> cloned into pWSK29	This study

Table 2.2 plasmids *cont*

pECs1581/pECs1581 (LI)	<i>ecs1581</i> amplified from strain TUV93-0 and cloned into pWSK29 (representative of Lineage I EHEC O157:H7 strain variants)	This study
pTW14539 (LI/II)	EHEC <i>ecs1581</i> orthologue (<i>ecsp_1496</i>) amplified from strain TW14539 and cloned into pWSK29 (representative of Lineage I/II and II O157:H7 strain variants)	This study
p96788 (LII)	EHEC <i>ecs1581</i> orthologue amplified from strain 96788 (un-sequenced) and cloned into pWSK29 (representative of Lineage II and LI/II O157:H7 strain variants)	This study
pC1493	<i>c1493</i> cloned into pWSK29	This study
pECED1_1787	<i>eced1_1787</i> cloned into pWSK29	This study
pQseC	<i>qseC</i> cloned into pWSK29	This study
pECED1_1787C20R	<i>eced1_1787</i> cloned into pWSK29 with a Cys to Arg substitution at position 20	This study
pECED1_1787C20R+N35K	<i>eced1_1787</i> cloned into pWSK29 with a Cys to Arg and Asn to Lys substitution at positions 20 and 35 respectively	This study

Table 2.2 plasmids *cont*

pECs1581R20C	<i>ecs1581</i> cloned into pWSK29 with an Arg to Cys substitution at position 20	This study
pECs1581R20C20N	<i>ecs1581R20C</i> cloned into pWSK29 with an Arg to Asn substitution at position 20	This study
pECs1581K35N	<i>ecs1581</i> cloned into pWSK29 with a Lys to Asn substitution at position 35	This study
pECs1581R20C+D10A	<i>ecs1581</i> cloned into pWSK29 with an Arg to Cys and Asp to Ala substitution at positions 20 and 10 respectively	This study
pECs1581R20C+K35N	<i>ecs1581</i> cloned into pWSK29 with an Arg to Cys and Lys to Asn substitution at positions 20 and 35 respectively	This study
pECs1581R20C+Q27K	<i>ecs1581</i> cloned into pWSK29 with an Arg to Cys and Gln to Lys substitution at positions 20 and 27 respectively	This study
pECs1581R20C+H45R	<i>ecs1581</i> cloned into pWSK29 with an Arg to Cys and His to Arg substitution at positions 20 and 45 respectively	This study

Table 2.2 plasmids *cont*

pECs1581R20C+R61D	<i>ecs1581</i> cloned into pWSK29 with an Arg to Cys and Arg to Asp substitution at positions 20 and 61 respectively	This study
Recombinant protein expression		
pET28a-ECs1581	<i>ecs1581</i> cloned into expression vector pET28a (+)	This study
pET28a-ECED1_1787	<i>eced1_1787</i> cloned into pET28a (+)	This study
pET28a-ECED1_1787C20R	<i>eced1_1787C20R</i> cloned into pET28a (+)	This study
H-NS	Purified recombinant H-NS protein gifted from Prof. Charles Dorman and Prof. Jay Hinton	(Tupper <i>et al.</i> , 1994)
Transcriptional fusion		
p <i>LEE1-gfp</i>	Upstream region of <i>LEE1</i> cloned into GFP+ reporter pKC26 to create transcriptional fusion construct p <i>LEE1::gfp</i>	This study
Allelic exchange		
pIBECs1581-US	Upstream flanking region of <i>ecs1581</i> cloned into pIB073	This study
pIBECs1581-DS	Downstream flanking region of <i>ecs1581</i> cloned into pIB073	

Table 2.2 plasmids *cont*

pIBECs1581- Kan ^R	Kanamycin resistance cassette cloned in between <i>ecs1581</i> US and DS flanking regions on pIB073	This study
pIBC1493-US	Upstream flanking region of <i>c1493</i> cloned into pIB073	This study
pIBC1493-DS	Downstream flanking region of <i>c1493</i> cloned into pIB073	
pIBC1493-Kan ^R	Kanamycin resistance cassette cloned in between <i>c1493</i> US and DS flanking regions on pIB073	
pIBQseC-US	Upstream flanking region of <i>qseC</i> cloned into pIB073	This study
pIBQseC-DS	Downstream flanking region of <i>qseC</i> cloned into pIB073	
pIBQseC-Kan ^R	Kanamycin resistance cassette cloned in between <i>qseC</i> US and DS flanking regions on pIB073	

Table 2.3 Primers pairs used in this research

Purpose/name	Sequence	Reference
Plasmid complementation		
z1835-z1843.F	CGTCTAGACTGCAACCGATTATAACGGATGCTTAACG	This study
z1835-z1843.R	GGAAGCTTTTACGCATGACTGTCCGGTGTTACTTCAG	
z1841.F	CGTCTAGAGAGGTGCATTAATGAAGG	This study
z1841.R	CGAAGCTTTAGCCTTCACGATGGTAG	
ecs1581.F	CGTCTAGACTTATCAGGTCACGCTATCG	This study
ecs1581.R	CGAAGCTTATCACCTCTGTGTCTTTGTG	
c1493.F	CGTCTAGATGAGTTGCAGGAAATAGAGA	This study
c1493.R	CGAAGCTT CGTCATGTTCCCTGCCATTG	
eced1_1787.F	CGAAGCTTTTACCTCTGTGTCTTTGTG	This study
eced1_1787.R	CGTCTAGAGCTTATCAGGTCACGCTATC	
TW14539/96788.F	CGTCTAGAAGAACAGCACACCGGAGAAT	This study
TW14539/96788.R	CGAAGCTTATGTGCCCAGGCATCACTAC	
ler.F	AAAAACTGCAGGTATCATATAGCATCATATAGTG	This study
ler.R	AAAAAGTCGACTCATGTAAATATTTTTCAGCGG	

Table 2.3 primers *cont*

qseC.F	CG TCTAGACATGTGCATCATCTGCGACG	This study
qseC.R	CG AAGCTTAGTACGTCACAGAGCGGTAC	
Site directed mutagenesis & plasmid complementation		
eced1_1787C20R.F	CGAAGCTTTTACCTCTGTGTCTTTGTGTTTTATTCCGCGTGGCT GTTTTTTTACGTTTCGGTGCCAGTTTCATTTCTTACCCCTGAATACG	This study
eced1_1787C20R.R	CGTCTAGACTATCGGGGAATAAAAATAATGAAACTTAAATAT CCTGGCTTAACTGCCAGCGGCAAACTCGCACTAAATTCATG CGCGGG	
eced1_1787C20R+N35K.F	CGAAGCTTCAGCTATGACCATGTATTACG	This study
eced1_1787C20R+N35K.nestR	TTCCTCCACGCCCTTAATCAT	
eced1_1787C20R+N35K.nestF	ATGATTAAGGGCGTGGAGGAA	
eced1_1787C20R+N35K.R	CGTCTAGACTGTGTCTTTGTGTTTTATTC	

Table 2.3 primers *cont*

ecs1581R20C.F	CGTCTAGACTATCGGGGAATAAAAATAATGAAACTTAAATAT CCTGGCTTAACTGACAGCGGCAAACTCGCACTAAATTCATG TGCGGG	This study
ecs1581R20C.R	CGAAGCTTATCACCTCTGTGTCTTTGTGTTTTATTCCGCGTGGC TGTTTTTTTACGTTCCGGTGCCAGTTTCATTTTTTACCCCTGAATAC	
ecs1581R20C20N.F	CGTCTAGACGCTATGACCATGATTACGC	This study
ecs1581R20C20N.nestR	GTAAATATCCCCGTTTCATGAA	
ecs1581R20C20N.nestF	TTCATGAACGGGGATATTTAC	
ecs1581R20C20N.R	CGAAGCTTATCACCTCTGTGTCTTTGTG	
ecs1581K35N.F	CGTCTAGACTTATCAGGTCACGCTATCG	This study
ecs1581K35N.nestR	TTCCTCCACGCCATTAATCAT	
ecs1581K35N.nestF	ATGATTAATGGCGTGGAGGAA	
ecs1581K35N.R	CGAAGCTTATCACCTCTGTGTCTTTGTG	

Table 2.3 primers *cont*

ecs1581R20C+D10A.F	CGTCTAGACTATCGGGGAATAAAAATAATGAAACTTAAATAT CCTGGCTTAACTGCCAGCGGC	This study
ecs1581R20C+D10A.R	CGAAGCTTCACCTCTGTGTCTTTGTGTTTTATTCCGCGTGGCTG TTTTTTTACGTTCCGGTGC	
ecs1581R20C+K35N.F	CGTCTAGACGCTATGACCATGATTACGC	This study
ecs1581R20C+K35N.nestR	TTCCTCCACGCCATTAATCAT	
ecs1581R20C+K35N.nestF	ATGATTAATGGCGTGGAGGAA	
ecs1581R20C+K35N.R	CGAAGCTTATCACCTCTGTGTCTTTGTG	
ecs1581R20C+H45R.F	CGTCTAGACGCTATGACCATGATTACGC	This study
ecs1581R20C+H45R.nestR	GTAGCCTTCACGGCGGTAAGT	
ecs1581R20C+H45R.nestF	ACTTACCGCCGTGAAGGCTAC	
ecs1581R20C+H45R.R	CGAAGCTTATCACCTCTGTGTCTTTGTG	

Table 2.3 primers *cont*

ecs1581R20C+R61D.F	CGTCTAGACGCTATGACCATGATTACGC	This study
ecs1581R20C+R61D.nestR	AGAAAAATCACGGTCGAACTG	
ecs1581R20C+R61D.nestF	CAGTTCGACCGTGATTTTTCT	
ecs1581R20C+R61D.R	CGAAGCTTATCACCTCTGTGTCTTTGTG	
ecs1581R20C+Q27K.F	CGTCTAGACGCTATGACCATGATTACGC	This study
ecs1581R20C+Q27K.nestR	CGTGCCGCCGTACTTATCGCG	
ecs1581R20C+Q27K.nestF	CGCGATAAGTACGGCGGCACG	
ecs1581R20C+Q27K.R	CGAAGCTTATCACCTCTGTGTCTTTGTG	
Recombinant protein expression		
ecs1581-pET28a.F	CGGGATCCATGAAACTTAAATATCCTGG	This study
ecs1581-pET28a.R	GGGTCTGACTACGTTTCGGTGCCAGTTTCA	
eced1_1787-pET28a.F	AAGTCGACTTACGTTTCGGTGCCAGTTTCA	This study
eced1_1787-pET28a.R	AAGGATCCATGAAACTTAAATATCCTGGC	

Table 2.3 primers *cont*

eced1_1787C20R-pET28a.F	AAGTCGACTTACGTCGGTGCCAGTTTCA	This study
eced1_1787C20R-pET28a.F	AAGGATCCATGAAACTTAAATATCCTGGC	
EMSA		
EMSA-LEE1.F	CATGCTTTAATATTTTAAGCTATTAGCGAC	This study
EMSA-LEE1.R	GCAATGAGATCTATCTTATAAAGAGAAACG	
EMSA-gapA.F	ACATTAACACCAACTGGCAAAATTTTGTCC	This study
EMSA-gapA.R	CATATATTCCACCAGCTATTTGTTAGTG	
EMSA-hns.F	CTCGTAAACACGACTAATACAGAAGACTG	This study
EMSA-hns.R	GTAGTAATCTCAAACCTTATATTGGGGTGGT	
RT-PCR		
RT-ler.F	CTGCGAGAGCAGGAAGTTCA	This study
RT-ler.R	AGGCACATTAGTATATCCCAGCTC	
RT-16S.F	ATTGACGTTACCCGCAGAAG	This study
RT-16S.R	CGCTTTACGCCAGTAATTC	

Table 2.3 primers *cont*
Screening and sequencing

pWSK29-screen.F	CAGTCACGACGTTGTAAAAC	This study
pWSK29-screen.R	CGTATGTTGTGTGGAATTGT	
pIB073-screen.F	CCTGTCCTACGAGTTGCATG	This study
pIB073-screen.R	GACTCCTGCATTAGGAAGCA	
pET28a-screen.F	CTTTGTTAGCAGCCGGATCTC	This study
pET28a-screen.R	CATCATCATCATCACAGCAGC	
pKC26-screen.F	GATGCAATTTCTATGCGCACC	This study
pKC26-screen.R	GTGAAGGTGATGCTACATACG	

Chapter 3

Phenotypic screening of O-island deletions

The genome of EHEC O157:H7 strain EDL933 contains 177 regions of DNA, termed OIs, which are absent from non-pathogenic *E. coli* strain K-12. Many of these islands contain sequences related to bacterial viruses and are believed to be horizontally acquired (Perna *et al.*, 2001). Many OIs have been suggested to be associated with virulence determinants but there is considerable uncertainty of gene function and their importance to the organism. This chapter describes the phenotypic characterisation of a large collection of OI deletions in EHEC O157:H7 strain TUV93-0 to gain a broad overview of what these horizontally acquired elements may actually be doing for this important pathogen. The hypothesis was that some of the OI deletion strains would differ when compared phenotypically to the parent strain.

3.1 Growth analysis

Bacterial generation time (g) was used as a marker to gauge the metabolic activity of EHEC strain TUV93-0 and a selection of sixteen isogenic OI mutants cultured in M9 minimal media at 37°C with agitation (200 rpm). Strains Δ OI-14-15, Δ OI-36, Δ OI-80 and Δ OI-148C all had doubling times (g = 89.0, 87.5, 87.3 and 88.9 minutes respectively) equivalent to the parent strain (g = 87.7 minutes) (Figure 3.1). Strains Δ OI-50, Δ OI-52 and Δ OI-79 grew somewhat faster than the

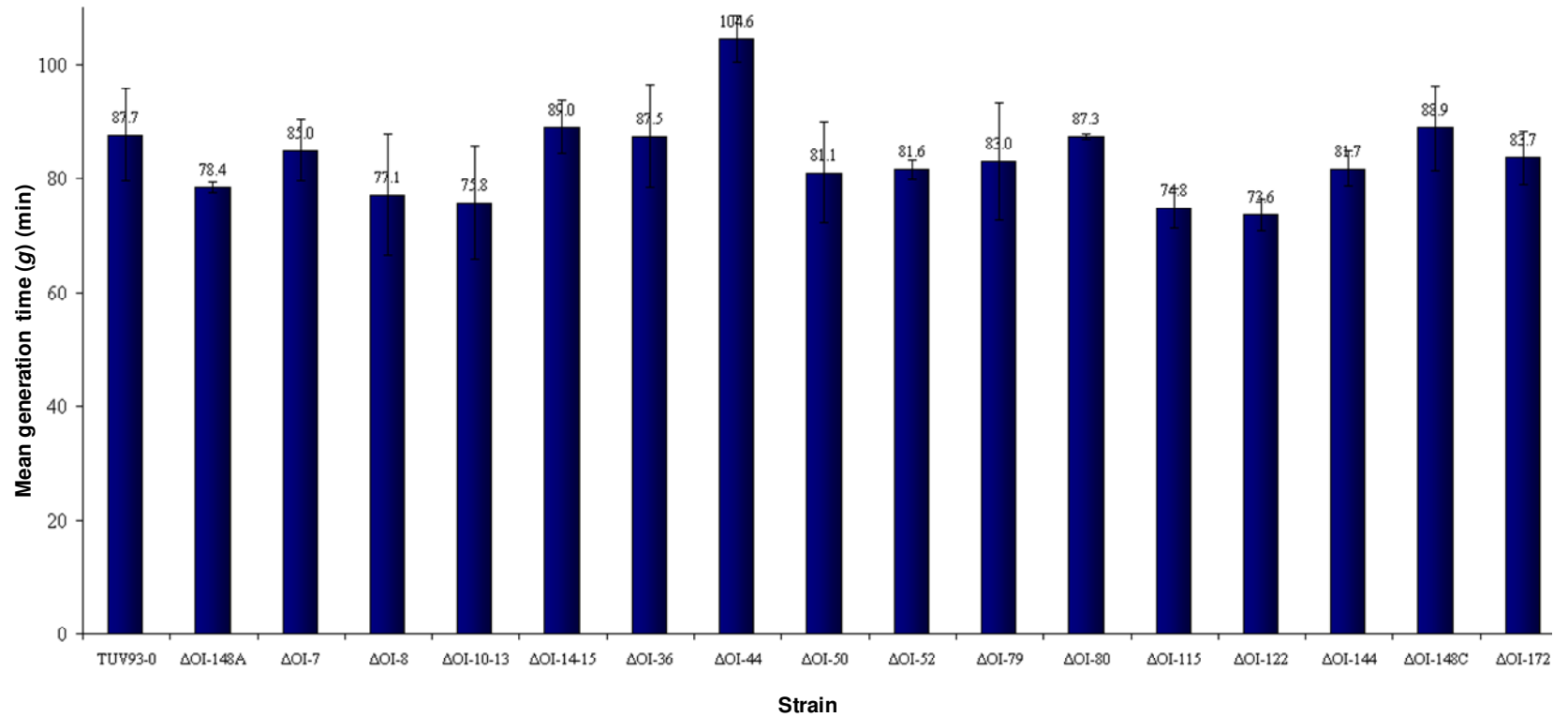


Figure 3.1. Growth phenotypes for EHEC O157:H7 strain TUV93-0 and a selection of sixteen isogenic OI deletion cultured in M9 minimal media at 37°C with shaking (200 rpm). Strain names are annotated as shown. The mean generation time for each OI-deletion was calculated as described in Chapter 2. Values represent the mean of two batches (n=2) for each mutant and the mean of 16 batches (n=16) for the parental strain.

parent with mean generation times ranging from between 81.0 to 83.0 minutes (Figure 3.1). Strains Δ OI-7, Δ OI-148A, Δ OI-8, Δ OI-10-13, Δ OI-115 and Δ OI-122 grew with shorter mean generation times ranging from between 73.6 minutes for Δ OI-122, the fastest replication rate of all the mutants tested, up to 78.4 minutes for Δ OI-148A (Figure 3.1). Strain Δ OI-44 had the slowest growth rate of all the mutants tested, with a difference in mean generation time of 16.9 minutes in comparison with the parent strain (Figure 3.1). These initial observations suggest that OI-122 somehow reduces the strains growth rate while OI-44 increases the replication rate of EHEC O157:H7 under the nutrient restricted growth conditions used in this study.

3.2 Acid survival

Acid resistance is perceived to be an important property for EHEC O157:H7 which enables survival under extremely acidic conditions such as those encountered in the mammalian GI tract. The possibility that OIs may in some way contribute to this important survival phenotype and potentially to other stress responses was therefore investigated. Twelve OI deletions and the parent strain were screened for their ability to survive acid challenge in LB broth pH 2.0. Samples were taken at the start of incubation ($t=0$) and 2 h post acid

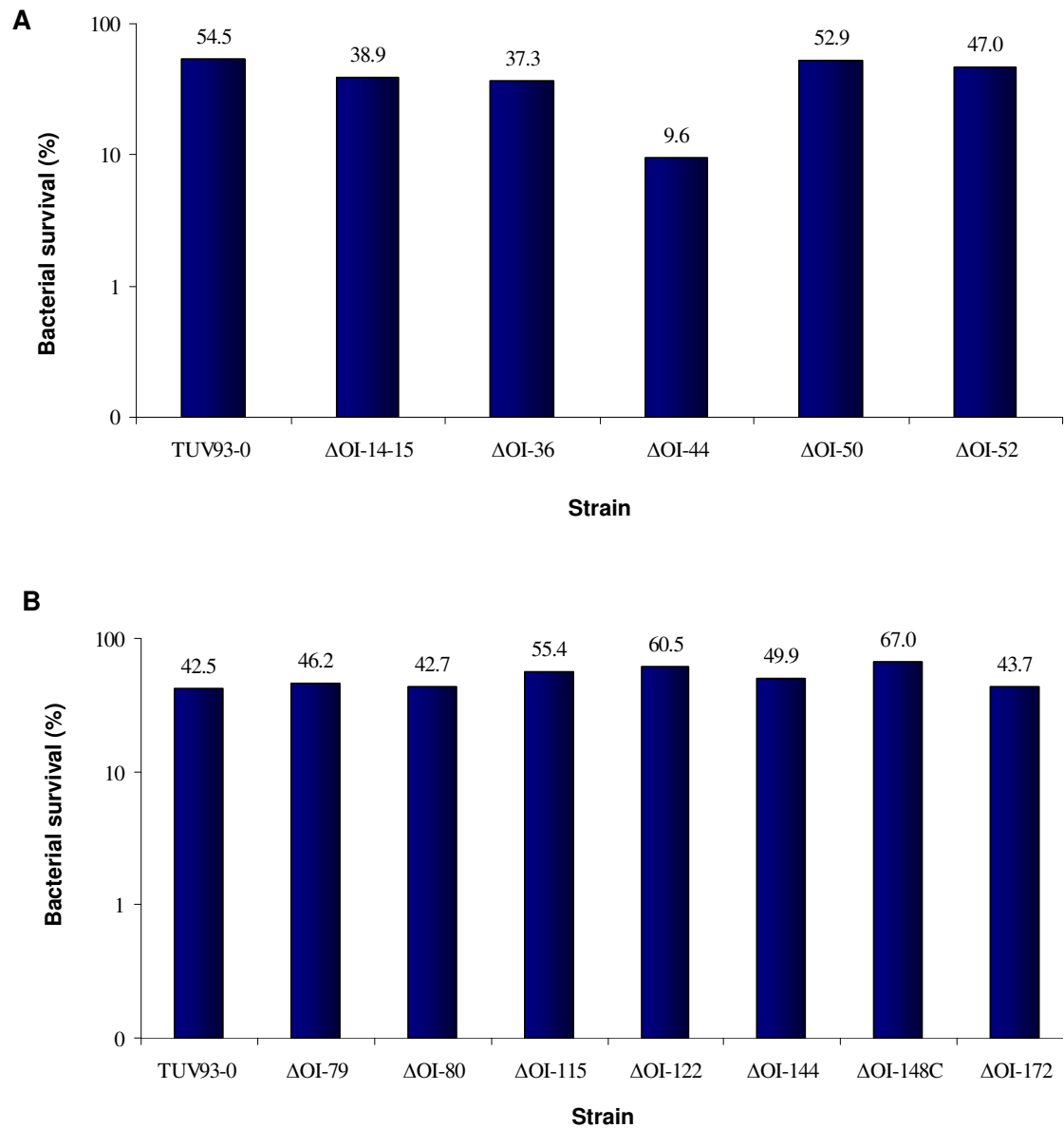


Figure 3.2. Acid survival phenotypes for EHEC O157:H7 strain TUV93-0 and a selection of twelve isogenic OI deletions growth in LB broth pH 2.0 for 2 h. (A-B) Challenge strains are annotated as shown and were assayed as described in Chapter 2. The number of colony forming units per millilitre (cfu/ml) for each strain was calculated post acid-challenge, and then compared with those from plated dilutions of the original culture to determine the % survival of each strain.

challenge ($t=2$) to allow the percentage survival of each strain to be calculated (cfu/ml) following plating out of the cultures onto LB agar plates and incubation at 37°C for 16 h. With the exception of Δ OI-44, all of the deletions demonstrated survival rates on a par with the parent strain post acid challenge (Figure 3.2A-B). Δ OI-44 displayed a fivefold reduction in survival compared to TUV93-0 (Figure 3.2A), and somewhat interestingly, this mutant also had a reduced growth rate in minimal M9 medium (Figure 3.1). Taking these two phenotypes together for this one mutant, these data indicate that OI-44 may be an important 'fitness' island for *E. coli* O157:H7.

3.3 Motility

A total of forty OI deletions were screened for motility by colony inoculation into soft agar and incubation at 37°C for 16 h. Although the majority of the OI deletions tested were as motile as the parent strain, deletion strains Δ OI-14-15 and Δ OI-52 were significantly less motile (Figure 3.3); suggesting that these islands may positively regulate cell motility. Likewise, several OI deletions had enhanced motility in comparison with the parent strain, for example Δ OI-44 and Δ OI-51; indicating that these islands may have a repressive effect on motility (Figure 3.3). Again, Δ OI-44 was phenotypically different from its parent strain,

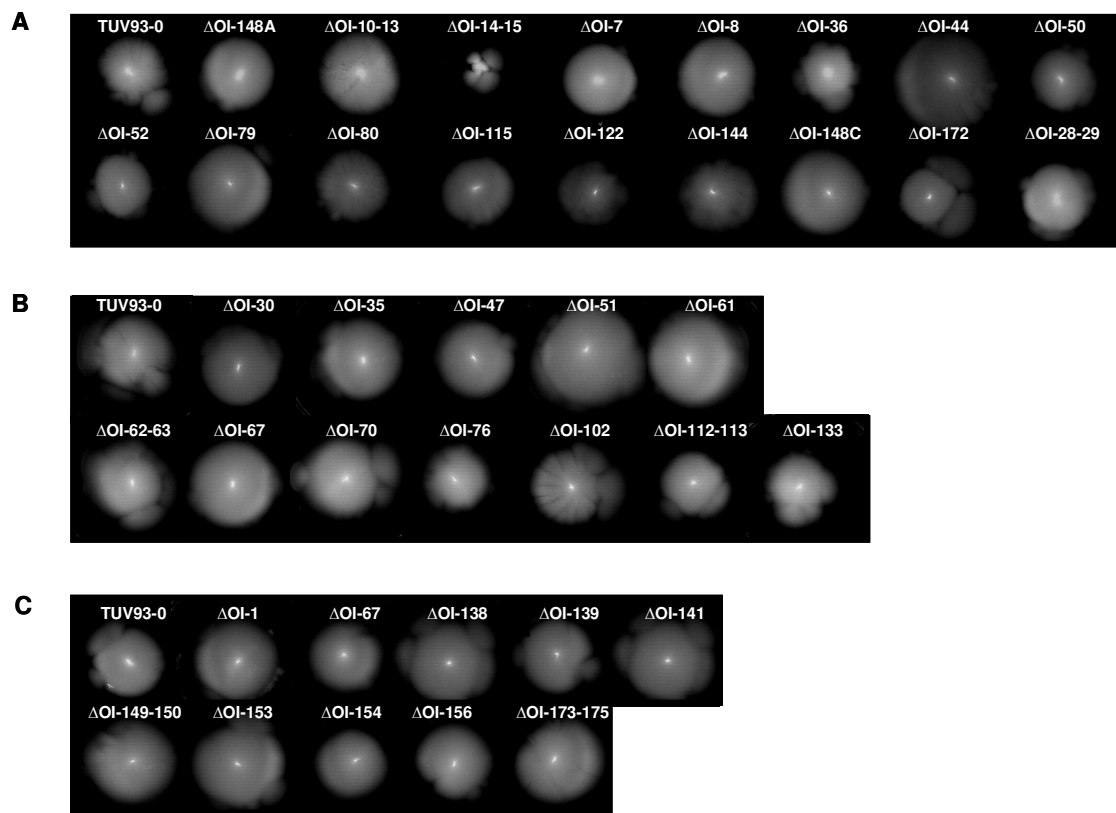


Figure 3.3. Analysis of motility in EHEC O157:H7 strain TUV93-0 and a selection of thirty nine isogenic OI deletions. (A-C) Test strains are annotated as shown. Strains were screened for motility as described in Chapter 2 and each experiment was carried out on at least three separate occasions.

suggesting this island has the potential to affect major phenotypes for O157:H7; although these remain to be thoroughly investigated. In relation to OI-14-15, this deletion had by far the clearest difference in motility compared with the parent strain (Figure 3.3), indicating that this island may have the capacity to either directly or indirectly affect bacterial motility in EHEC O157:H7.

3.4 Analysis of T3S culture supernatant proteins

Type III secretion (T3S) is an essential virulence determinant for EHEC that is required for successful colonisation of the bovine host (Naylor *et al.*, 2004). To investigate the impact of OIs on T3S levels, twenty four OI deletions were cultured in MEM--HEPES at 37°C with shaking (200 rpm). Supernatants were filtered, TCA precipitated and analysed by SDS-PAGE and colloidal blue staining. Initial screening identified a subset of OIs with the capacity to either repress or activate T3S. For example, Δ OI-47, Δ OI-76 (Figure 3.4A) and Δ OI-141 (Figure 3.4B) deletions all had levels of T3S above that of the wild type parent, suggesting repression by these islands; whilst Δ OI-51, Δ OI-70 (Figure 4.3A) and Δ OI-133 (Figure 4.3B) deletions had reduced levels of T3S, suggesting activation by these islands. Of note was the OI-51 deletion, as this was attenuated for T3S,

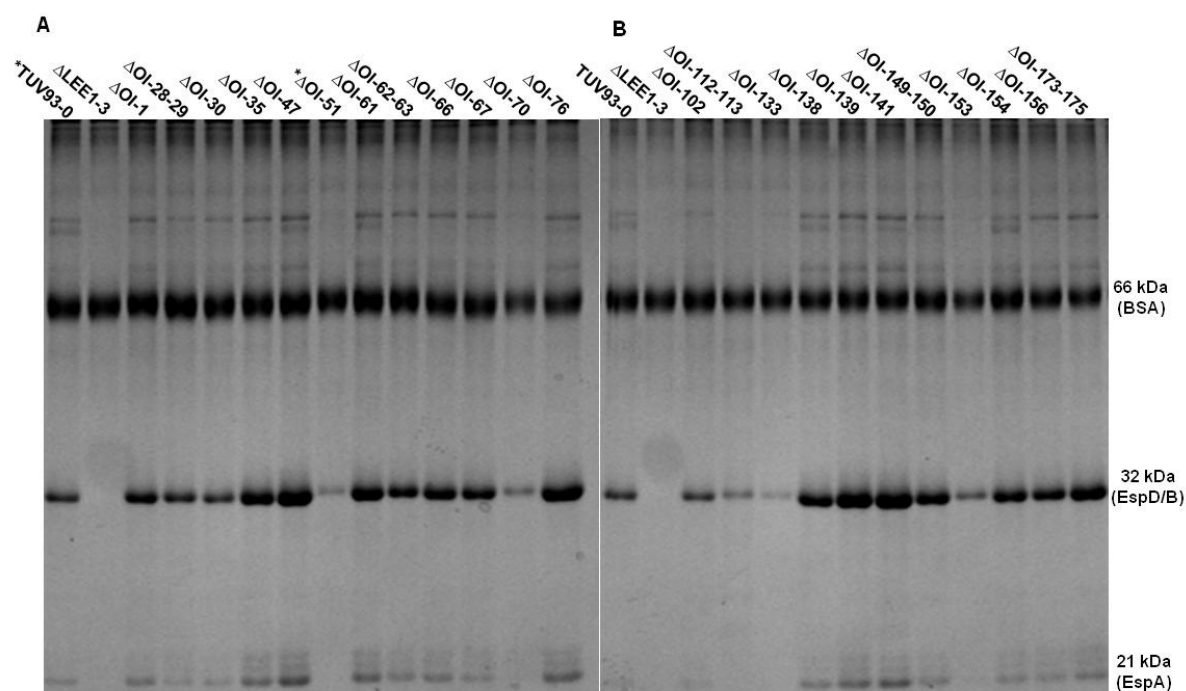


Figure 3.4 Screening of EHEC O157:H7 OI deletions for altered levels of T3S
(A-B) SDS-PAGE gels showing culture supernatant secretion profiles for EHEC strain TUV93-0 and a selection of twenty four isogenic deletion strains. Strain names are annotated as shown. Parent strain TUV93-0 acted a positive control for T3S and a T3S system mutant (Δ LEE1-3) provided a negative control for secretion. The translocon protein bands are indicated, EspB/D and EspA as well as Bovine serum albumin (BSA) which was added as a loading control to help rule out the precipitation process as a source of variation and to act as a co-precipitant. The wild type parent and OI-51 deletion strain that was taken forward for further analysis are marked with an asterisk for ease of viewing. Proteins were prepared and analysed as described in Chapter 2.

and also displayed enhanced motility compared with the parent strain. Although speculative at this stage, OI-51 may have the capacity to reciprocally control motility and T3S; regulatory cross-talk that has already been observed in EHEC. OI-51 is therefore identified as a potentially important OI controlling T3S and motility and was taken forward for further investigation.

3.5 Discussion

EHEC O157:H7 is a serious human pathogen that can cause bloody diarrhoea, haemorrhagic colitis and haemolytic uremic syndrome. EHEC pathogenesis is a diverse and often complex process involving the co-ordinated expression of virulence factors and transcriptional regulators that enable the bacteria to survive and cause disease, such as can be found on horizontally acquired genomic elements, for example on plasmids and in the genome of bacteriophage, that are transferred between bacteria.

In the chromosome of EHEC O157:H7 strain EDL933 a large number of multigenic regions, or OIs, have been identified that include both fully functional and cryptic prophages (Perna *et al.*, 2001); and studies on EHEC have demonstrated that the genotypic diversity observed amongst different strains of

EHEC O157:H7 is largely attributable to these bacteriophage-associated fragments (Ohnishi *et al.*, 2002; Zhang *et al.*, 2007). Given that these regions can to some extent be considered accessory to the core genome, their presence in EHEC O157:H7 indicates selection in relation to the specific 'lifestyle' of the organism; for example, EHEC colonisation of the bovine gastrointestinal tract at the terminal rectum (Naylor *et al.*, 2003). It is now evident that a number of the larger OIs are cryptic lambdoid phage encoding effector proteins that are secreted through the T3S and which contribute to EHEC O157 colonisation through effects on adherence and immune modulation. Indeed, A number of OIs have already been partially characterised, for example, OI-7 (putative macrophage toxin), OI-10-13 (EaeH putative adhesin), OI-28 (RTX-toxin-like exoprotein), OI-50 (SopA like protein; Psr regulators), OI-71 (NleF, NleH1-2, and NleA), OI-79 (EspFu effector protein), OI-80 (putative invasion), OI-115 (ETTA2), OI-122 (cytotoxin), OI-144 (putative adhesin) and OI-148 (LEE PAI) (Campellone *et al.*, 2003; Perna *et al.*, 2001).

Although not all of the OIs identified in the genome of EDL933 will have a role in virulence, those that do not may still have important biochemical and physiological functions. Further characterisation of these OIs is therefore crucial

to enable a clearer understanding of the mechanisms utilised by this important zoonotic pathogen that help it colonise at the bovine terminal rectum; and that help the bacterium survive *in vivo* on passage to this colonisation site; including its persistence and survival in the environment and transmission to other animals and humans. Therefore, an initial aim of this research was to phenotypically screen a large selection of OI deletions in EHEC strain TUV93-0; to try and identify novel prophage regions that confer important phenotypes on the bacterium which may help it to survive and colonise in its mammalian host. It was envisaged that at least a few OI deletions would be sufficiently different from their parent strain so as to provide a subset of targets regions that could be systematically analysed through cloning and complementation and targeted mutagenesis.

Of interest, an OI-44 deletion strain was shown to have a growth rate considerably slower (~17 min) than its wild type parent following culture in minimal M9 medium (Figure 3.1). Furthermore, this mutant was phenotypically more motile and less resistant to acid stress. OI-44 is a 9.3 kb partial cryptic prophage, originally designated as CP-933M (Campellone *et al.*, 2004; Perna *et al.*, 2001) that is composed of 15 open reading frames. Many of these are of

unknown function although this deletion has been reported to form A/E lesions on HeLa cells at wild type levels (Campellon *et al.*, 2004). Preliminary evidence here may support this prophage as being important for EHEC O157:H7 survival, although this would clearly require further investigation. Also of interest was an OI-122 deletion which had a growth rate considerably quicker (~ 14 min) than its parent. OI-122 is a 21.2 kb genomic island composed of 25 open reading frames, some of which share significant homology to virulence genes found in *S. enterica* (*pagC*) and *S. flexneri* (enterotoxin 2 or *sen*) (Karmali *et al.*, 2003; Perna *et al.*, 2001); and recently a number of T3S effector proteins have been identified on OI-122 (*ent*/*EspL2*, *NleB*, and *NleE*), although this island was not screened for T3S in this study; however, Campellone *et al* (2004) have previously demonstrated that an OI-122 deletion strain can form wild type levels of A/E lesion on cells. The pattern of *Nle* genes encoded in OI-71 and OI-122 have also recently been closely associated with certain serotypes of EHEC and EPEC that are associated with severe illness and outbreaks in children (Bugarel *et al.*, 2010; Bugarel *et al.*, 2011). As the loss of OI-122 from the genome of TUV93-0 appeared to increase growth rate, this may suggest that OI-122 has the potential to suppress bacterial metabolism and/or cell division. Again however, the full significance of these findings remains to be fully established.

Acid resistance enables organism such as *E. coli* O157:H7 to survive gastric acidity (pH <3.0) and volatile fatty acids produced as a result of fermentation in the intestine (Castanie-Cornet *et al.*, 1999). The ability to resist these acid stresses is believed to be necessary for *E. coli* to colonise its mammalian host and the low infectious dose associated with pathogenic *E. coli* O157:H7 is attributed to its acid resistant nature (Castanie-Cornet *et al.*, 1999). With this in mind, an acid shock assay was carried out to assess the acid resistance phenotype for a number of OI deletions following acid challenge at pH 2.0 for 2 hours. Initial findings were somewhat disappointing in that most of the deletions assayed displayed no significant reduction in survival compared with the wild type parent; although one deletion (Δ OI-44) did show a small reduction (5 fold) in survival. The significance of this result however was uncertain due to a number of issues relating to this assay; such as harsh pH followed by mechanical stress on plating out. Moreover, this assay was also extremely labour intensive, and so on this basis, it was decided not to pursue this island.

For many bacteria, flagella play an important role in bacterial adaptation to environmental conditions and have also been associated with virulence (Yona-Nadler *et al.*, 2003). A selection of sixteen TUV93-0 derived OI deletions were

therefore examined for differences in this phenotype. Of particular interest was an OI-14-15 deletion as this strain was significantly less motile than its parent strain. OI-14-15 is a 16.5 kb genomic island composed of 9 genes and 11 open reading frames, and is believed to contain a putative adhesion (Campellone *et al.*, 2004; Wells *et al.*, 2008). The flagellar system of *E. coli* is known to be encoded by over 40 genes organised into several co-regulated operons (Yona-Nadler *et al.*, 2003). Over a dozen of these operons have been discovered and are classified into three hierarchical classes (classes 1, 2 and 3), which together form the flagellar regulon (Iyoda *et al.*, 2006). It may be that there is cross-talk between OI-14-15 and the flagellar regulon in TUV93-0, although this island was over-looked for further investigation in place of another OI target, OI-51, which was shown to affect motility and T3S; the latter being an extremely important virulence mechanism required for EHEC colonisation in ruminants (Naylor *et al.*, 2004).

The delivery of virulence factors directly into host cells to interfere with and alter host processes is a crucial step in bacterial virulence. In *E. coli* O157:H7, this is achieved through the LEE encoded T3SS which the bacteria utilises to inject a number of T3SS effector proteins into the cytosol of host cells, resulting

in cytoskeleton actin rearrangement and pedestal formation (A/E lesion). The expression of LEE and T3SS proteins however is a complex and coordinated process responding to many environmental signals, such as temperature, pH, osmolarity, growth phase and quorum sensing (Laaberki *et al.*, 2006). In this study, secreted protein profiles for a number of OI mutants were examined and compared with their parent strain. MEM-HEPES and DMEM were used as the growth medium of choice as studies have shown that levels of type III secretion are strain and medium dependant, with MEM-HEPES and DMEM being known promote the secretion of T3SS proteins such as EspA, EspB and EspD (Roe *et al.*, 2003). Major differences in secretion levels allowed a number of mutants to be classified as either high secretors or low secretors in comparison with the parent. It would have been exciting to pursue a number of these OIs, but due to time restrictions, it was decided to take forward only one of these mutants for further investigation. Given that an OI-51 mutant was attenuated for T3S and appeared more motile than its parent strain, the prospect that this prophage region may encode a potentially new regulator able to coordinate T3S and motility in EHEC swayed the decision to study this island. The genotypic and phenotypic characterisation of this prophage region is now detailed in Chapter 4.

Chapter 4

Genotypic and phenotypic characterisation of OI-51

T3S is essential for EHEC O157:H7 colonisation in ruminants (Naylor *et al.*, 2005) and previous analyses here identified an OI-51 deletion as having markedly reduced levels of T3S compared with its wild type parent. This chapter goes on to describe the genotypic and phenotypic characterisation of this potentially significant region in EHEC O157:H7.

4.1 *In vivo* analysis of OI-51

To assess the potential importance of OI-51 for colonisation, an OI-51 deletion and parent strain were competed in an established ovine colonisation model (Wales *et al.*, 2001). The OI-51 deletion was marked with a chloramphenicol resistance cassette and the parent strain by selection for resistance to Naladixic acid to allow direct plating and enumeration of the two strains from faeces (Figure 4.1). The strains were competed in 6 animals, only one of which failed to be properly colonised (< 20 CFU/gram faeces from the daily counts). For the remaining five animals the estimated relative total levels of the mutant excreted from day 5 onwards, to remove inoculum effects, were estimated (Figure 4.1). A significant reduction was demonstrated for the mutant ($p=0.004$) indicating that OI-51 is indeed important for colonisation and persistence in the ruminant GI tract.

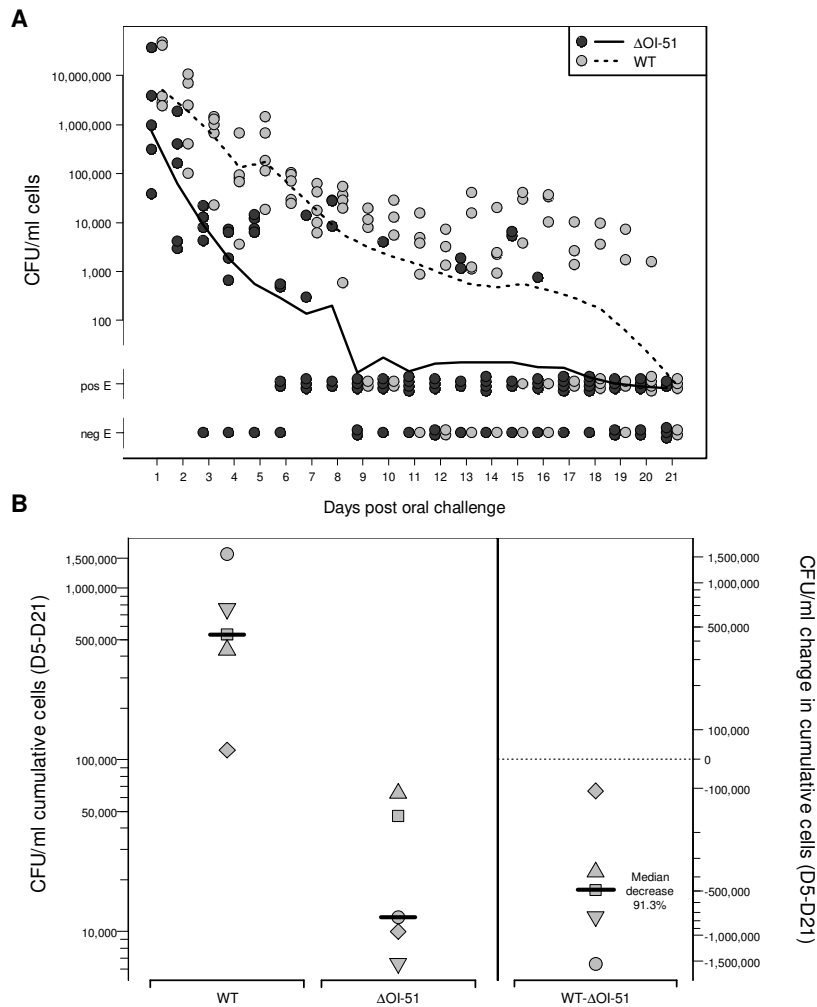


Figure 4.1. Contribution of OI-51 to EHEC colonisation in sheep. Six animals were orally dosed with both wild type (TUV93-0) and OI-51 deletion strains as described in Chapter 2. **(A)** Levels of both strains were determined daily from faecal samples. 'PosE' are samples that were positive for the strain following broth enrichment; 'NegE' samples were negative following broth enrichment. **(B)** The cumulative shedding levels for each animal and strain were estimated for colonised animals from day 5 onwards to avoid inoculum effects. One animal was excluded from the analysis that was not properly colonised by either strain (<20 CFU per gram of faeces). The different symbols represent individual animals and allow direct comparison of the two strains in each animal. The right hand panel of the graph shows the percentage mean decrease of the mutant strain by comparison with the WT. The difference in median cumulative shedding levels is statistically significant ($p=0.004$). All animal work was carried out by the VLA Weybridge and the associated statistical analyses were carried out with help from Dr Darren Shaw.

4.2 OI-51 sequence analysis

OI-51 is a 14.93 Kb cryptic prophage designated as CP-933C in sequenced EHEC O157:H7 strain EDL933, and Sp7 in closely related EHEC O157:H7 strain Sakai (Figure 4.2). Bioinformatic analyses of this genomic region show it to be an unusual and highly degraded prophage comprised mainly of P4 phage remnants. The majority of open reading frames annotated in OI-51 are hypothetical although several share features with known proteins; including a P4 integrase (similar to CP4-like integrase and integrase used for 933L and LEE PAI); a P4-like excisionase (Xis); a replication gene similar to the P4 α gene; a putative DNA binding protein similar to P4 ORF88; a putative single stranded DNA binding protein; a putative transcriptional activator, PchE, that shares homology to the plasmid encoded regulator C (PerC) locus in enteropathogenic *E. coli* (EPEC); and phage structural genes. Although many morphogenesis genes are absent in OI-51, evidence suggests this prophage can still be packaged and replicated, and it is hypothesised that OI-51 may be complemented for these defects by procuring the relevant proteins from other prophages in the O157:H7 chromosome; through prophage-prophage interactions (Asadulghani *et al.*, 2009)

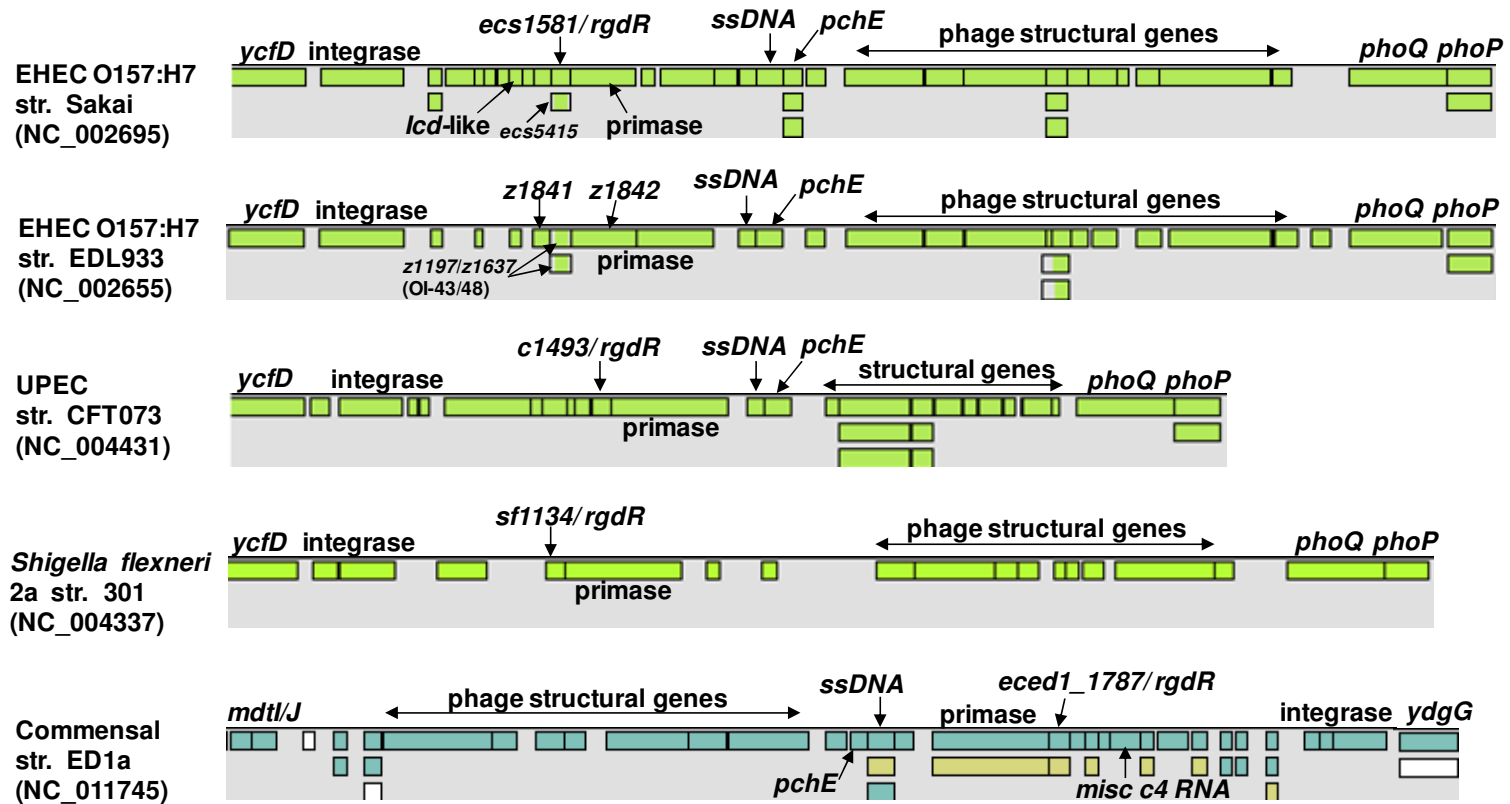


Figure 4.2. Schematic representation of the genomic organisation of OI-51 and similar prophages in a subset of *E. coli* and *Shigella* reference strains (EHEC O157:H7 strains EDL933 and Sakai, UPEC strain CFT073, *Shigella flexneri* (2a) strain 301 and commensal strain ED1a). Putative characterised genes are annotated as shown. Sequence related alleles identified elsewhere in the genome are represented by shaded boxes in a vertical column.

4.3 OI-51 cloning and complementation

In order to genotypically investigate Δ OI-51s reduced T3S phenotype further, OI-51 genomic fragments were PCR amplified (Figure 4.3A) from strain TUV93-0 and cloned into the commercial cloning vector TOPO (Invitrogen), to create expression plasmids pTZ1835-Z1843 and pTZ1850-Z1857; and then these were assessed for their effects on T3S levels following transformation into the OI-51 deletion strain (Figure 4.3B). Unfortunately, attempts to clone the middle portion of OI-51 (harbouring genes *z1844-z1849*) were unsuccessful at this point. The potential significance of this prophage region will be discussed later as this harbours the previously mentioned *pchE* allele, a PerC like that is can be found in certain EPEC strains(*perA-C*); termed *pchA-E* in EHEC, that can regulate LEE gene expression (Abe *et al.*, 2008; Bustamante *et al.*, 2011; Iyoda and Watanabe, 2004; Porter *et al.*, 2005). Following culture, expression construct pTZ1835-43 was able to rescue T3S in the OI-51 mutant by restoring T3S to the levels demonstrated in the parent strain, indicating that a positive regulator of the LEE was encoded somewhere on this fragment (Figure 4.3B). By comparison, expression construct pTZ1850-57 had no discernable effect on T3S levels and was therefore regarded as a good internal cloning control (Figure 4.3B). Subsequent cloning of fragment *z1835-z1843* into the low copy number vector

pWSK29 (pWZ1835-43) also rescued T3S when expressed in the OI-51 deletion strain (Figure 4.3C); reducing the likelihood that the regulation was due to issues associated with the use of a high copy number vector.

4.4 Analysis of cloned OI-51 fragment *z1835-z1843*

Previous analyses identified pTZ1835-43/pWZ1835-43 (harbouring OI-51 genes *z1835-z1843* on high and low copy number vectors respectively) as a 5.8 Kb cloned fragment that is able to stimulate T3S in EHEC O157:H7 (Figure 4.3B). This finding was informative as it enabled an original ~15 Kb (OI-51) DNA region of interest to be narrowed down to 5.8 Kb. To now narrow this region down even further, a series of restriction endonuclease cut-backs were made in pWZ1835-43, creating expression plasmids pXbaI+SphI, pSphI+HpaI, pHpaI+BstZ171, pHpaI+Bsu361 and pBsu361+HindII (Figure 4.4A-E); and then these were screened for their effects on T3S following Klenow-fill in, re-ligation and then transformation back into the OI-51 mutant. The rationale for this being that any loss in activity (i.e. the construct is no longer able to rescue T3S) is indicative of the regulator having been removed; thus localising the regulator to that fragment. T3S levels were significantly reduced in only one of the constructs; the OI-51 deletion harbouring pHpaI+Bsu361 (missing orfs *z1840*,

z1841 and *z1842*) (Figure 4.4F). As expression plasmids pHpaI+BstZ17I (missing orf *z1840*) and pBsu361+HindIII (missing half of orf *z1842* and all of *z1843*) were still able to stimulate T3S in the OI-51 deletion (Figure 4.4F), this ruled out all of the genes harboured on these fragments as being the regulator (s); leaving only orf *z1841* and an intergenic region between *z1841* and *z1842* (Figure 4.4D). When *z1841* was provided *in trans* on low copy number vector pWSK29 (pZ1841), this construct was unable to rescue T3S levels in the OI-51 mutant (Figure 4.4H). This unexpected result prompted closer inspection of the intergenic region between *z1841* and *z1842* (Figure 4.4I). In closely related EHEC O157:H7 strain Sakai, an open reading frame is annotated between these two open reading frames, designated as *ecs1581* (Figure 4.4I). Successful PCR amplification of *ecs1581* using EDL933 derived strain TUV93-0 as a genomic template confirmed the presence of *ecs1581* in this strain (data not shown). Consequently, *ecs1581* was cloned into pWSK29 (pECs1581) and analysed for T3S in the OI-51 mutant.

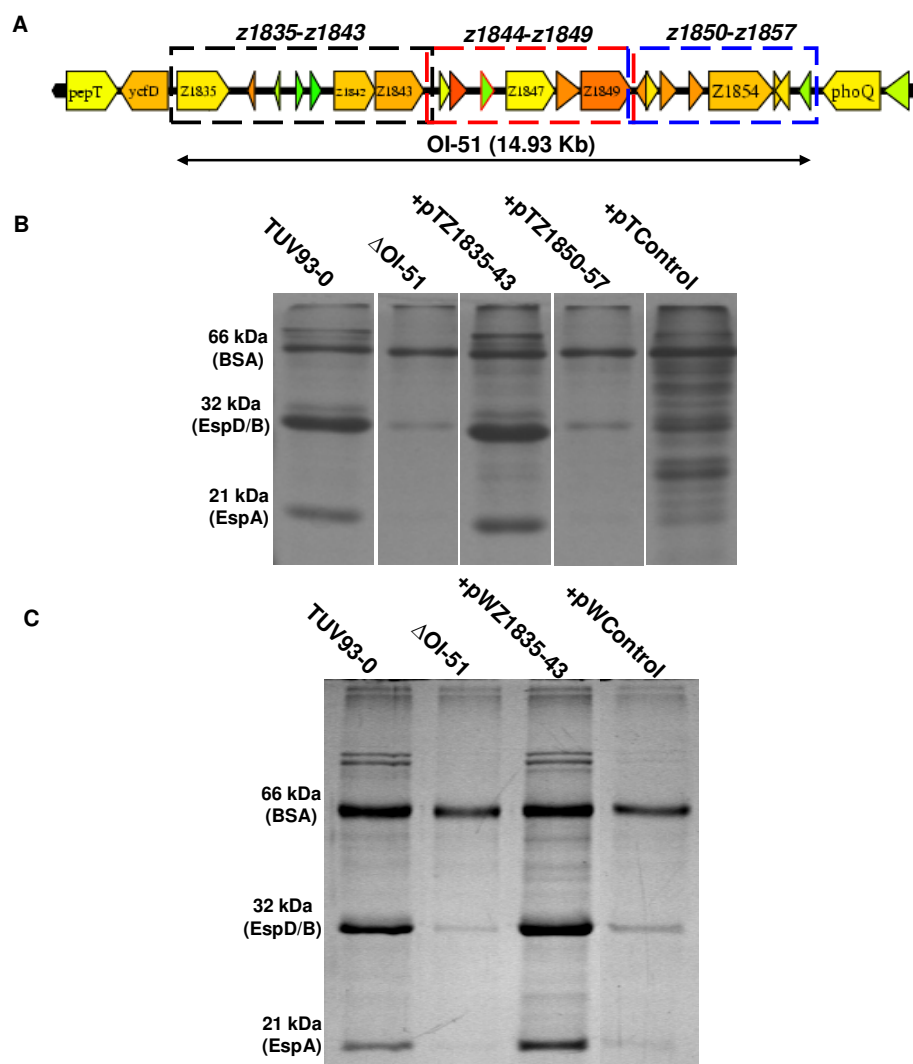


Figure 4.3. Analysis of T3S levels in an OI-51 deletion *in trans* expressing cloned OI-51 fragments. (A) OI-51 genomic fragments (*z1835-z1843* and *z1850-z1857*) were PCR amplified from parent strain TUV93-0 and cloned into TOPO to create expression constructs pTZ1835-Z1843 and pZ1850-Z1857. (B) Analysis of T3S levels in an OI-51 deletion expressing pTZ1835-Z1843 or pTZ1850-Z1857. Empty vector alone (pTControl) acted as a control (C) Analysis of T3S levels in an OI-51 deletion expressing pWZ1835-43 (T3S rescue fragment *z1835-z1843* re-cloned into low copy number cloning vector pWSK29). Proteins were prepared and analysed as described in Chapter 2.

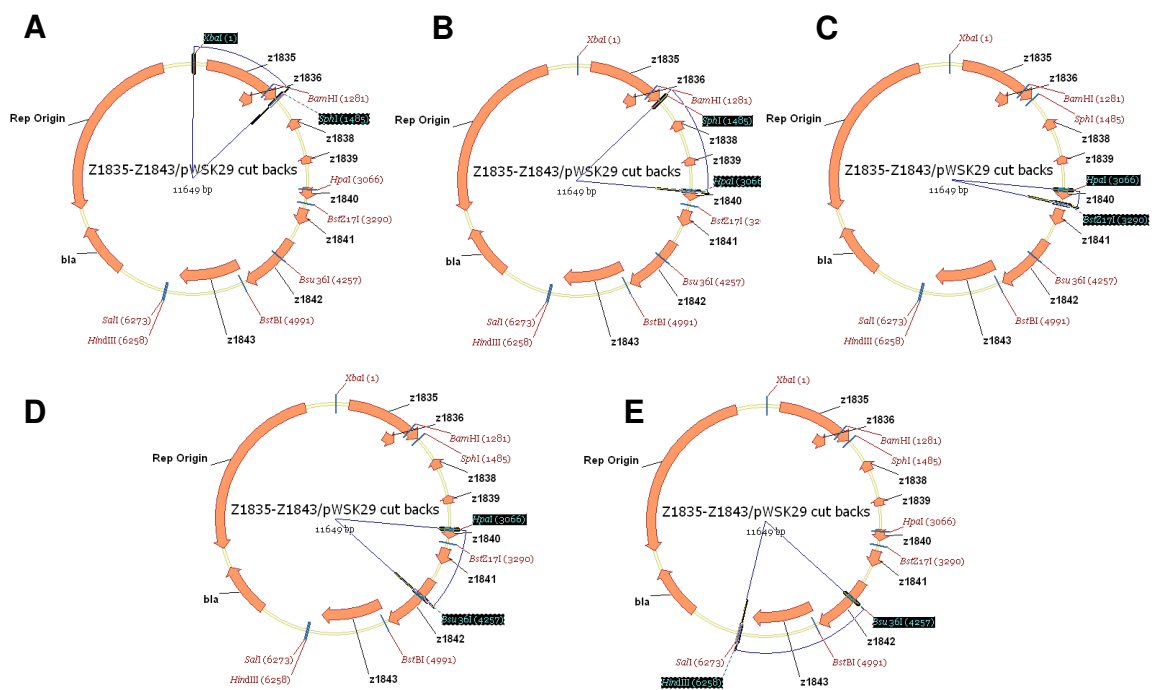


Figure 4.4. Construction of pWZ1835-43 cut-back derivatives and the analysis of their effects on T3S levels. (A) Schematic representation of the cut-backs made in pWZ1835-43. Allelic fragments were cleaved from T3S rescue construct pWZ1835-43 using the restriction enzymes detailed in panel 4.4G. Resulting constructs (pXbaI+SphI, pSphI+HpaI, pHpaI+BstZ171, pHpaI+Bsu36I or pBsu36I+HindIII) were Klenow filled, re-ligated and then transformed into an OI-51 deletion to assess their effects on T3S levels (Figure 4.4F).

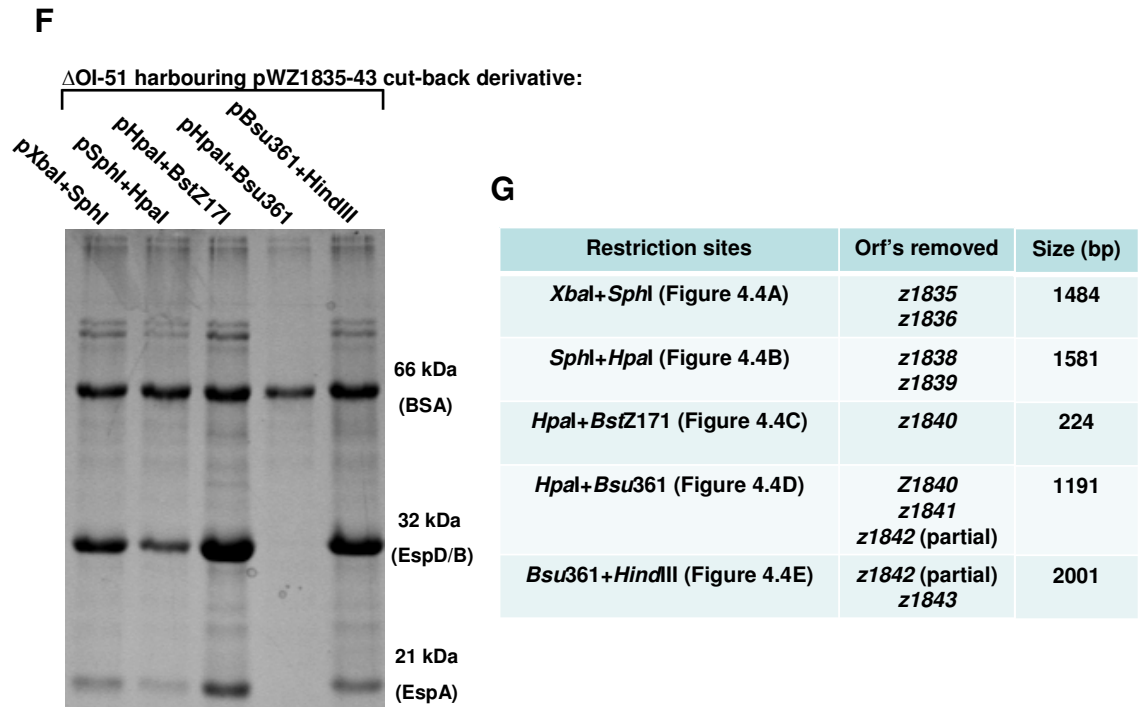


Figure 4.4 cont. An 1191 bp cloned fragment from OI-51 can stimulate T3S in EHEC O157:H7 (F) SDS-PAGE gel showing T3S levels in an OI-51 deletion expressing cut-back derivatives of pWZ1835-43 (annotated as shown); where construct *pHpaI*&*Bsu361* (no longer harbouring alleles *z1840*, *z1841* and the first half of *z1842* (partial)) has lost the capacity to activate T3S, indicating that a putative regulator of T3S is located somewhere in this region. Protein samples were prepared and analysed as described in Chapter 2. **(G)** Summary panel detailing the restriction sites that were used to digest-out alleles from T3S rescue fragment pWZ1835-43. The corresponding digest sizes are shown.

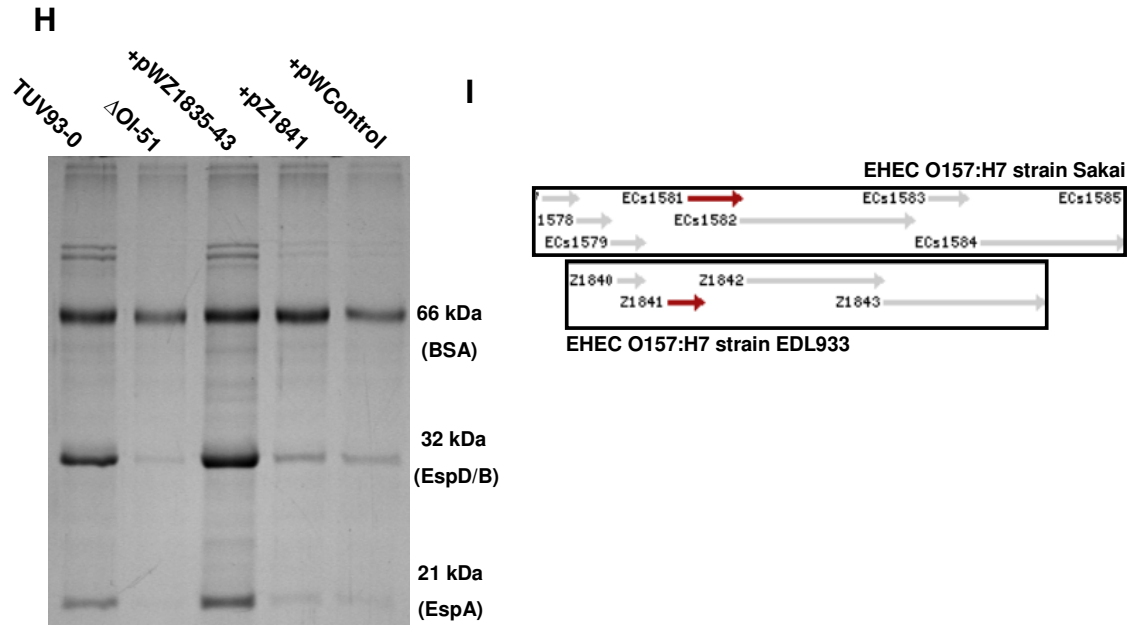


Figure 4.4 cont. Z1841 is unable to rescue T3S levels in an OI-51 deletion. (H) *z1841* was cloned into low copy number plasmid pWSK29 (pZ1841) and analysed for its effects on T3S levels in an OI-51 deletion; to ascertain if this protein is capable of rescuing T3S levels. Empty vector alone (pWControl) acted as a negative control and the original T3S rescue construct pWZ1835-43 acted as a positive control. Z1841 was unable to complement the OI-51 deletion for T3S indicating it is not a T3S regulator. Protein samples were prepared and analysed as described in Chapter 2. (I) Schematic of the genetic organisation of open reading frames in the region *z1840-z1843* in EHEC O157:H7 strains EDL933 and Sakai; where several differences in annotation can be seen between the two strains. Of pertinence is an open reading frame annotated only in the Sakai sequence (*ecs1581*) but between *z1841* and *z1842* in EDL933; precisely where a putative T3S regulator is predicted to be.

4.5 Identification of *ecs1581* as a positive regulator of T3S

To establish *ecs1581* as a positive regulator of T3S encoded on OI-51, *ecs1581* was cloned (pECs1581) and induced in an OI-51 deletion strain and then analysed for T3S. ECs1581 strongly stimulated secretion in this background, bringing levels back to beyond wild type and equivalent to the positive control fragment pWZ1835-43. An *ecs1581* mutant was then constructed by inserting a kanamycin resistance cassette into the open reading frame of this allele in TUV93-0. This strain was then analysed for T3S along with an *ecs1581* deletion harbouring ECs1581 *in trans* (pECs1581) (Figure 4.5B). T3S levels were significantly reduced in the *ecs1581* mutant. Further, these reduced levels of T3S could be rescued when *ecs1581* was introduced in low copy number plasmid (pECs1581); in both the defined *ecs1581* deletion and OI-51 mutant (Figure 4.5B). *ecs1581* was therefore identified as a putative new regulator capable of stimulating T3S in EHEC O157:H7.

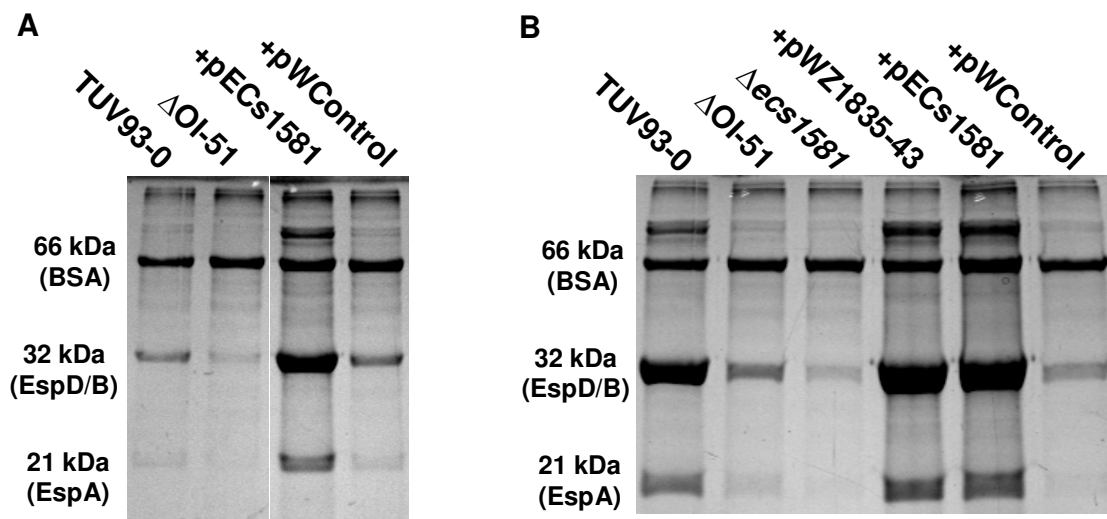


Fig. 4.5 Identification of *ecs1581* as a positive regulator of T3S. (A) SDS-PAGE gel showing T3S complementation in an OI-51 deletion expressing ECs1581 on low copy number vector pWSK29 (pECs1581). Empty vector alone (pWControl) acted as a control. (B) SDS-PAGE gel showing significantly reduced levels of T3S in an *ecs1581* deletion strain and then the rescue of this phenotype when ECs1581 is provided *in trans*. Construct pZ1835-43 was included for reference and as a positive control for secretion. The empty vector alone (pWControl) acted as a negative control. Protein samples were prepared and analysed as described in Chapter 2.

4.6 Discussion

Genomic variation in *E. coli* has been shown to be predominately associated with horizontally-acquired genomic islands and phage-related regions that often harbour virulence-associated genes. During this research, a number of EHEC prophage deletions (OIs) were examined for their effect on T3S (an essential virulence factor for EHEC which is required for successful colonisation of the ruminant host) in order to identify horizontally-acquired regions that cross-talk with this crucial colonisation system (Chapter 3). Of the twenty four OI mutants studied in TUV93-0 (excluding the LEE1-3 deletion), five had lower levels of T3S, including OI-51. OI-51 is a 14.93 kb region in EDL933/TUV93-0 and 15.46 kb (Sp7) in the sequenced EHEC O157:H7 strain Sakai (Hayashi *et al.*, 2001; Perna *et al.*, 2001). Variations of this island are present in all EHEC O157:H7 strains as well as non-O157 EHEC strains, but also in specific UPEC, NMEC and commensal *E. coli* strains that do not contain a LEE-encoded T3S system. Deletion of OI-51 reduced persistence in a competitive ruminant colonisation experiment indicating the potential importance of this region. Subsequent cloning and complementation in an OI-51 deletion identified a small gene on OI-51, *ecs1581* that was able to stimulate T3S in EHEC. OI-51 in the sequenced *E. coli* O157:H7 EDL933 strain contains a predicted 22 open reading

frames with most indicating P4 ancestry (including a P4 integrase; a P4-like excisionase; a replication gene similar to the P4 α gene; and a putative DNA binding protein similar to P4 *orf88*), although a putative single stranded DNA binding protein (*ssDNA*) and a transcriptional activator (encoded by *pchE*) sharing homology to the PerC locus in EPEC can also be found on this island. The genome of EHEC O157:H7 contains five genes homologous to the PerC protein which is produced by certain EPEC strains (Abe *et al.*, 2008; Bustamante *et al.*, 2011; Iyoda and Watanabe, 2004; Porter *et al.*, 2005). Two of these *perC* homologue genes, *ecs1388* (*pchD*) and *ecs1588* (*pchE*) encode proteins with 25% and 39% sequence identity with respect to EPEC PerC. Whilst other *pch* (*pchA*-*C*) genes have been demonstrated to modulate the expression of LEE transcription units in EHEC O157:H7, deletion or over-expression of *ECs1388* or *ECs1588* have not; and so their function, if any, in gene regulation remains to be determined (Iyoda *et al.*, 2004; Porter *et al.*, 2005; Abe *et al.*, 2008; Yang *et al.*, 2009). It is therefore of interest that a new potential regulator of T3S (encoded by *ecs1581*) has now been discovered on OI-51 adjacent to *pchE*. *ECs1581* was able to up-regulate T3S in both a defined *ecs1581* and OI-51 deletion, indicating that *ECs1581* can exert its control over T3S in the absence of *pchE* and the other genes on this island. Interestingly, suppressive subtractive hybridisation

studies have demonstrated that the presence of OI-51 is more closely associated with human infection strains of EHEC (Steele *et al.*, 2007). Again, it is noteworthy that a novel positive regulator of T3S has now been indentified on OI-51. In addition, several studies have reported clear differences in expression levels of T3S proteins by *E. coli* O157:H7 (such as *stx* production and *LEE* gene expression) isolated from humans and from cattle (Roe *et al.*, 2003; Roe *et al.*, 2004; Steele *et al.*, 2007); and DNA microarray studies (Dowd and Ishizaki, 2006) have also demonstrated that lineage specific strains of O157:H7 have evolved distinct patterns of gene expression. As ECs1581 has the capacity to stimulate T3S, this regulator may have important implications for the persistence and survival of EHEC *in vivo* in both the ruminant and coincidental human host.

EHEC strain Sakai, a close relative of EDL933, contains 18 prophages (Sp1 to Sp18) and 6 prophage like elements (SpLE1 to SpLE6), amounting to 16% of the total genome (Hayashi *et al.*, 2001; Ohnishi *et al.*, 2001; Asadulghani *et al.*, 2009). Extensive characterisation of these prophage regions by Asadulghani and co-workers (2009) have indicated that most of these prophage (Sp3– Sp6, Sp8–Sp12, Sp14 and Sp15) retain features of lambdoid phages; where one (Sp13) contains features similar to P2 phage, one (Sp1) contains P4 features, and one (Sp18)

retains Mu features. The other four prophages, including OI-51/Sp7 (Sp1 Sp7, Sp16, and Sp17) are un-assigned to a particular phage family due to their chimeric and highly disrupted genomic backbones (Asadulghani *et al.*, 2009). In the case of OI-51/Sp7, this cryptic prophage lacks most morphogenesis genes, including repressor and anti-repressor genes, and the replication gene that is similar to the α gene of satellite phage P4 has also been disrupted in Sakai by multiple frame shift mutations (Asadulghani *et al.*, 2009). However, Sp7 has been shown to still be inducible. Thus, it has hypothesised that the replication of Sp7 may be facilitated by the procurement of replication proteins from other prophages in the chromosome, for example Sp13 (P2-like phage) or Sp2 (P4-like phage); through prophage-prophage interactions. However, these authors concur that some (or all) of the fragmented polypeptides of Sp7 may still contain some replication initiation activity (Asadulghani *et al.*, 2009).

In summary, this chapter of work has identified a region in EHEC (OI-51) that is important for EHEC colonisation in ruminants; and has identified ECs1581 as a new regulator encoded on OI-51 that has the capacity to stimulate T3S. The functional and mechanistic characterisation of this regulatory protein is described in Chapter 5.

Chapter 5

Functional characterisation of ECs1581/RgdR

Cloning and complementation of OI-51 in combination with targeted mutagenesis identified ECs1581 as a small and novel protein that is able to stimulate T3S in EHEC O157:H7. This chapter describes the further characterisation of ECs1581 together with a selection of sequence related variants that are present in other *E. coli* strains; including strains not harbouring a LEE-encoded T3SS.

5.1 Distribution of ECs1581 in *E. coli* strains

ECs1581 is predicted to be a small protein of 99 amino acids with a molecular mass of 11.704 kDa and a pI of 10.09. ECs1581 has no significant sequence similarity with any known protein although it does have a very short section of homology with the spacer region of the LEE encoded regulator (Ler) that bridges the known protein-protein interaction domain of Ler with its DNA binding domain (Figure 5.1) (Mellies *et al.*, 2008). However, there is no significant homology of ECs1581 with these functional regions of Ler (Figure 5.1) (Mellies *et al.*, 2008).

Sequence analysis of ECs1581 protein variants in different lineages of EHEC O157:H7 showed the core region of this protein (amino acids 20-68) to be highly

Ler	1	MRRLFIMNMENNSHTTSPYIQLIEQIAVLQQEAKRLREQ	VQSVIESIQK	50
ECs1581	0	-----		0
		Ler linker region		Ler/H-NS DNA binding motif
Ler	51	QITYYMITLQELGYTNVPDDGLARRNSSKGVY	RNEEGQ	100
		.: . .:. . . .:. . .:. . . .:		
ECs1581	1	-----	MKLKYPGLTDSGKTRTKFMRGDIYRDQYGGT	31
Ler	101	LKEALLNGMKKEDFLVKDTEEEIPLKNI	-----	129
		.: . . .:. . .:. . . .:		
ECs1581	32	-----	VMIKGVEERRVTYHREGYEYDCVMPVYQFRDFSLVQ	68
Ler	129	-----		129
ECs1581	69	AAPRSKPTSREKARANIQEIKMLNVFRGKK		99

Figure 5.1. Sequence comparison of ECs1581 with LEE-encoded regulator (Ler). Alignment of ECs1581 and Ler amino acid sequences showing slight homology between the two proteins at the N-terminus of ECs1581 and the linker/oligomerisation region of Ler. Ler's established DNA binding motif and linker region are annotated as shown (Mellies *et al.*, 2008).

conserved; with distinct lineage specific amino acid substitutions at the N- and C-termini (Figure 5.2A). Interestingly, this sequence divergence is specific to lineage I (LI) EHEC O157:H7 by comparison with lineage II (LII) and lineage I/II (LI/II) O157:H7 (Figure 5.2A). Orthologues of ECs1581 can be found in sequenced non-O157 EHEC strains (O103:H2, O26:H11, and O111:H-); in EPEC strain O55:H7; newborn meningitis *E. coli* (NMEC) strain IHE3034; enteroaggregative *E. coli* (EAEC) strain O42; atypical EPEC strain E110019; uropathogenic *E. coli* (UPEC) strains CFT073 (O6:K2:H1) and IAI39 (O7:K1); *Shigella spp* and commensal *E. coli* strain ED1a (O81); all with >77% homology over a region of 99 amino acids (i.e. full length ECs1581) (Table 5.1). Multiple divergent copies of *ecs1581* are also present in the chromosomes of some strains harbouring this regulator, examples of which are also detailed in Table 5.1. Although some variants are truncated (63 amino acids), they do contain a putative protein interaction module, PB1, which can be found in a range of signalling proteins in eukaryotes and plants (Moscat *et al.*, 2006; Sumimoto *et al.*, 2007; Terasawa *et al.*, 2001). Likewise, many of the full length variants share a putative N-myristylation motif (GLTD/ASG; amino acids 7-12), which appears to have diverged in the EHEC O157:H7 LI variants of ECs1581 with an alanine to aspartic acid substitution at amino acid position 10 (A10D) (GLTDSG/LI and

A

Strain	Lineage	20	40	60	80	100							
Sakai	I	MKLKYPGLTD	SGKTRTKFMR	GDIYRDQYGG	TVMIKGVEER	RVTYHREGYE	YDCVMPVYQF	RRDFS	LVQAA	PRSKPTSREK	ARANIQEIKK	MLNVFRGKK*	100
EDL933	I	100
EC869	IIS...A	..N.HP..TR.....T	.HNV...NARKL.T	.V.G.....	100
96788	IIS...A	..N.HP..TR.....T	.HNV...NARKL.T	.V.G.....	100
TW14539	I/IIS...A	..N.HP..TR.....T	.HNV...NARKL.T	.V.G.....	100
EC4113	I/IIS...A	..N.HP..TR.....T	.HNV...NARKL.T	.V.G.....	100

B

UPEC C1493	MKLKYSGLTASGKTRTKFMRGDIYRDQYGGTVMIKGVEERRVTVRREGYEYDCVMPVYQF	60
EHEC ECSP_ 1496 (LI/II&LII)	MKLKYSGLTASGNTHPKFTIRGDIYRDQYGGTVMIKGVEERRVTVRREGYEYDCVMPVYQF	60
EHEC ECs1581 (LI)	MKLKYPGLTISGKTRTKFMRGDIYRIQYGGTVMIKGVEERRVTVRREGYEYDCVMPVYQF	60
ED1a ECED1_1787	MKLKYPGLTASGKTRTKFMRGDIYRIQYGGTVMIKGVEERRVTVRREGYEYDCVMPVYQF	60
	*****.*** **:.*:..** *****:*****:*****:*****:*****	
UPEC C1493	RRDFS	LVQTAPHNVSTSNARARANIQLKTMINGFRGKK 99
EHEC ECSP_ 1496 (LI/II&LII)	RRDFS	LVQTAPHNVPTSNARARANIQLKTMVNGFRGKK 99
EHEC ECs1581 (LI)	RRDFS	LVQAAPRSKPTSREKARANIQEIKKMLNVFRGKK 99
ED1a ECED1_1787	RRDFS	LVQAAPRSKPTSREKARANIQEIKKMLNVFRGKK 99
	*****:***. **:.*:..** *****:***:*** *****	

Figure 5.2. Sequence comparison of ECs1581 variants in *E. coli*. (A) Multiple alignment of the predicted amino acid sequences of ECs1581 in different lineages of EHEC O157:H7. Residue divergence specific to lineage I EHEC O157:H7 isolates is marked by a pink shaded box. (B) Multiple alignment of EHEC variants ECs1581 (EDL933/Sakai) and ECSP_1496 (TW14539) with commensal variant ECED1_1787 (ED1a) and UPEC variant C1493 (CFT073). Stars indicate conserved residues. Putative domains are marked by black boxes (putative RGD motif at amino acid positions 20-22 and a putative N-myristylation site at amino acid positions 7-12). Residue divergence between EHEC ECs1581 and commensal ECED1_1787 is outlined with a red box and residue divergence between EHEC LI and LI/II&LII variants of ECs1581 is indicated by a black underscore.

GLTASG/LII and LI/II) (Figure 5.2B). ECs1581 also contains a tri-peptide arginine-glycine-aspartate (RGD) motif (amino acids 20-22, Figure 5.2B) which may be important for mediating protein-protein interactions based on previously established functions of this motif (D'Souza *et al.*, 1991).

5.2 Analysis of ECs1581 orthologous protein C1493

Given the diversity of *E. coli* strains that harbour sequence related variants of ECs1581, its prevalence and divergence amongst different EHEC lineages, and the fact that not all strains with this regulator harbour a LEE-encoded T3S system, the functionality of a subset of these protein variants were investigated by testing their effects on T3S and the related but more ubiquitous flagella system.

UPEC CFT073 variant C1493 has 83% sequence similarity to ECs1581 (Figure 5.2B) and is also encoded on a cryptic prophage (CP073-3) that is inserted in an identical location on the chromosome as CP-933C (OI-51/Sp7) (Mobley *et al.*, 1990; Perna *et al.*, 2001; Welch *et al.*, 2002). *c1493* was cloned into the low copy number plasmid pWSK29 to create expression plasmid pC1493 and this was assessed for its effects on T3S levels following induced expression in an EHEC

ecs1581 mutant. Expression of pC1493 rescued T3S demonstrating that the variants are functionally interchangeable, despite UPEC lineages not harbouring a LEE encoded T3S system (Figure 5.3A).

To investigate motility, a *c1493* mutant was constructed in UPEC strain CFT073 and examined in soft agar following in *trans* expression of *c1493* and *ecs1581*; in the reciprocal manner motility was examined in a EHEC *ecs1581* mutant strain with in *trans* expression of the two variants. Although only subtle differences in motility were observed for wild type and mutant constructs in both EHEC and UPEC, the *in trans* and presumably higher level expression of the two variants down regulated motility, a phenotype that was not observed for the strains expressing the empty vector (pWControl) (Figure 5.3B).

5.3 Analysis of ECs1581 activity in a *grlA* mutant

In EHEC, GrlA is an established regulatory protein encoded by the LEE that positively regulates *ler* transcription and LEE gene expression; and negatively regulates motility in a feedback loop with the repressor GrlR (Iyoda *et al.*, 2006). To rule out the possibility that in EHEC ECs1581 may be acting on T3S and

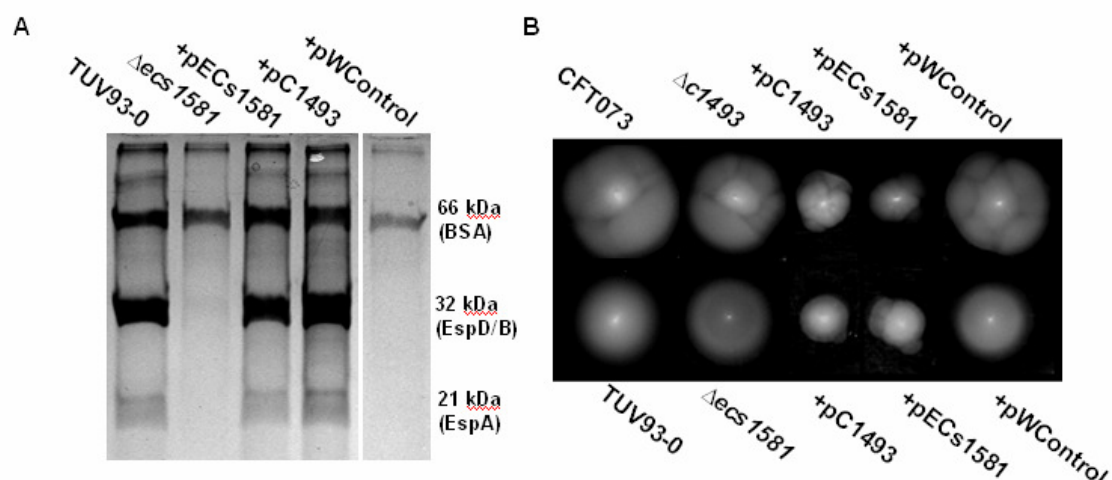


Figure 5.3. ECs1581 and orthologous protein C1493 can activate T3S and down-regulate motility in EHEC and UPEC. (A) SDS-PAGE gel showing stimulation of T3S in an EHEC *ecs1581* deletion *in trans* expressing EHEC derived ECs1581 (pECs1581) or UPEC derived C1493 (pC1493). Empty vector (pWControl) alone in the *ecs1581* mutant acted as a negative control strain. (B) Analysis of motility repression in an EHEC *ecs1581* and UPEC *c1493* deletion background *in trans* expressing pECs1581, pC1493 or pWControl. Strains were analysed for T3S and motility as described in Chapter 2.

motility via *grlA*, T3S and motility was assessed in a *grlA* mutant with induced expression of *ecs1581* from a low copy number plasmid (pECs1581). As expected, T3S levels were reduced in the *grlA* mutant compared with the parent strain (Figure 5.4A). Induced expression of *ecs1581* in the *grlA* deletion was then able to increase T3S levels in this strain (Figure 5.4A); indicating that ECs1581 activation of T3S is mediated independently of GrlA. Moreover, induced expression of ECs1581 was enough to repress motility in this strain (Figure 5.4A-B).

5.4 Analysis of ECs1581 orthologous protein ECED1_1787

Commensal strain ED1a (a close relative of UPEC strain CFT073; phylogenetic group B2) also harbours an orthologue of ECs1581, ECED1_1787, that has 94% sequence similarity (Figure 5.2B) (Clermont *et al.*, 2008; Jaureguiy *et al.*, 2008; Touchon *et al.*, 2009). Despite having the highest known homology to ECs1581, even compared to other EHEC O157:H7 variants, commensal protein ECED1_1787 was unable to activate T3S or repress motility following induced expression (pECED1_1787) in an EHEC *ecs1581* mutant strain (Figure 5.5A). This was informative as there are only six amino acid differences between ECs1581 and ECED1_1787 (Figure 5.2B); and by the finding that C1493 has this

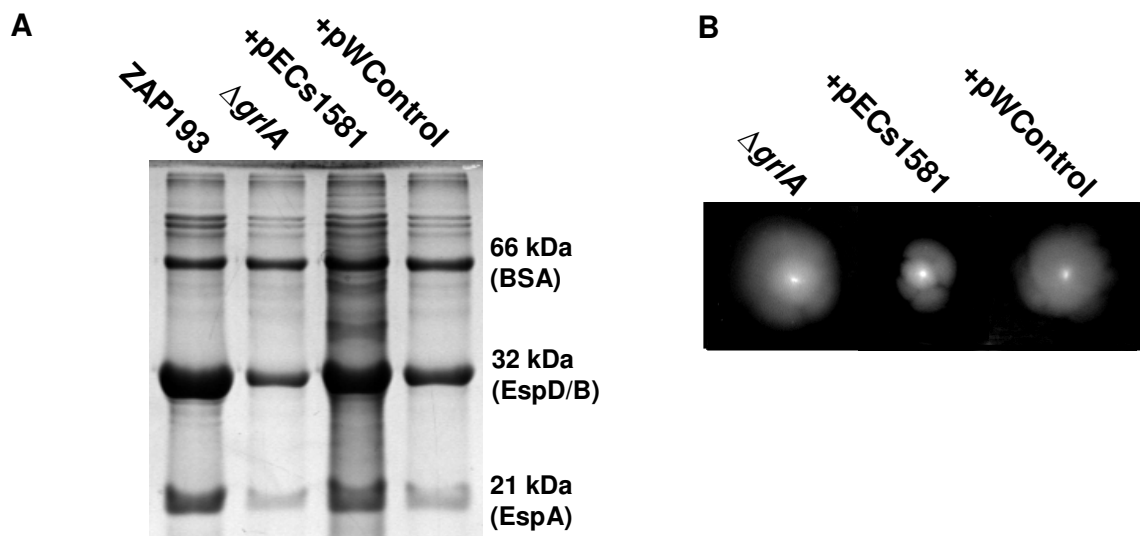


Figure 5.4. ECs1581 can stimulate T3S independently of GrIA. (A) SDS-PAGE gel showing T3S profile for EHEC strain ZAP193 and isogenic $\Delta grlA$::Tn5 mutant *in trans* expressing ECs1581 (pECs1581) or empty vector control (pWControl). (B) Motility plates showing repression in a *grlA* mutant expressing pECs1581. Culture supernatant proteins were prepared and analysed as described in Chapter 2.

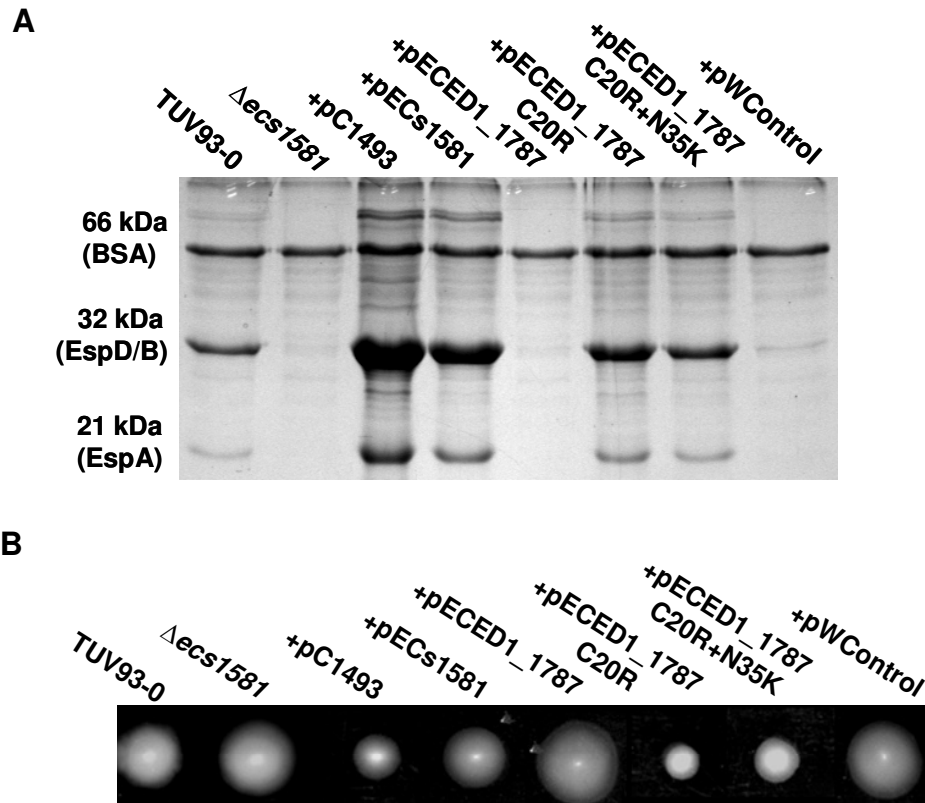


Figure 5.5. Effects of natural and engineered variants of ECs1581 on T3S and motility.

(**A**) SDS-PAGE gel showing T3S levels in an EHEC TUV93-0 *ecs1581* mutant *in trans* expressing wild type and engineered variants of commensal protein ECED1_1787 (pECED1_1787, pECED1_1787C20R and pECED1_1787N35K). pWControl in the *ecs1581* deletion was used as a negative control and wild type variants pECs1581 and pC1493 were included as a reference for high T3S levels. (**B**) Analysis of motility in an *ecs1581* mutant *in trans* expressing pECED1_1787, pECED1_1787C20R, pECED1_1787N35K or pWControl. Strains were analysed for T3S and motility as described in Chapter 2.

capacity despite its lower homology to ECs1581 (83% versus 94% respectively). The divergence in ECED1_1787 by comparison to ECs1581 maps to the following residues: D10A; R20C; Q27K; K35N; H45R and R61D (amino acids 1-99) (Figure 5.2B). Interestingly, two of these changes occur within predicted domains with a D10A substitution in the putative N-myristylation motif (amino acids 7-12), and an R20C substitution in the EHEC conserved RGD motif (amino acids 20-22) (Figure 5.2B). Like ECED1_1787, C1493 has a D10A but harbours an RGD motif as in ECs1581 (Figure 5.2B).

5.5 Analysis of natural and engineered variants of ECED1_1787 and ECs1581

Based on the functional differences demonstrated for the ECs1581 and ECED1_1787 proteins, a series of reciprocal substitutions in proteins ECs1581 and ECED1_1787 were constructed by site-directed mutagenesis and their effects on T3S and motility assessed in an EHEC *ecs1581* mutant. A single C20R mutation in protein ECED1_1787 (pECED1_1787C20R) was sufficient to enable this variant to now activate T3S and repress motility (Figure 5.5A). An ECED1_1787 C20R+N35K double mutation (pECED1_1787C20R+N35K) did not increase T3S levels over the single C20R change (Figure 5.5A). Strikingly, induced expression of ECED1_1787C20R severely repressed motility (Figure

5.5B). This shift in phenotype between wild type ECED1_1787 and engineered variant ECED1_1787C20R was intriguing as induced expression of ECED1_1787 in the same strain background increased motility (Figure 5.5B). A K35N substitution in ECs1581 (ECED1_1787N35) limited its stimulation of T3S in an EHEC *ecs1581* mutant when induced (pECs1581K35N) from a low copy number plasmid (Figure 5.5C). Surprisingly, an R20C change in the EHEC conserved RGD motif (ECED1_1787C20) slightly increased T3S levels following induced expression (pECs1581R20C) in this background, and this engineered variant was also more repressive for motility than wild type ECs1581 (Figure 5.5C/D). A further R20C20N mutation in this same RGD motif however did reduce T3S levels compared to wild type ECs1581 (Figure 5.5C). Taken together, these results highlight the functional importance of residue 20 and other targeted residue combinations for the control exhibited by this novel regulator.

5.5.1 Analysis of EHEC lineage specific variants of ECs1581

Sequence analyses of ECs1581 in EHEC O157:H7 have shown that two main variants of ECs1581 are present. These variants can be categorised according to the O157:H7 lineage from which they derive; where Lineage I variants of ECs1581 are distinct from lineage II and lineage I/II variants which share

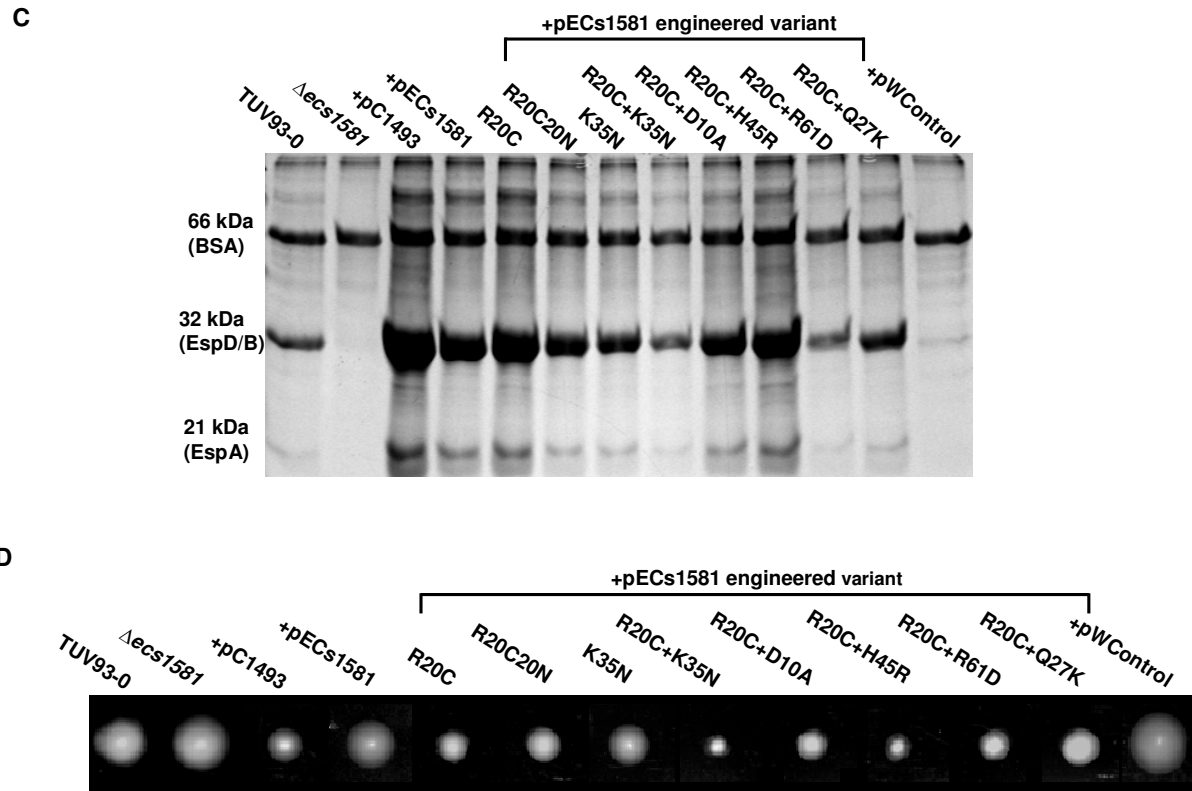


Figure 5.5 continued. (C) SDS-PAGE gel showing T3S levels in an *ecs1581* mutant expressing wild type and engineered variants of ECs1581 (pC1493, pECs1581, pECs1581R20C, pECs1581R20C20N, pECs1581K35N, pECs1581R20C+K35N, pECs1581R20C+D10A, pECs1581R20C+H45R, pECs1581R20C+R61D, pECs1581R20C+Q27K and pWControl). (D) Analysis of motility in the same strain set. The motility figures (B and D) are composite images from photographs of motility agar plates inoculated with the strain of interest. The same images are used for the wild types and controls in B and D. Strains were analysed for T3S and motility as described in Chapter 2.

identical protein coding sequences. It was therefore a possibility that these lineage II and lineage I/II variants may have differing capacities to stimulate T3S and repress motility compared with lineage I derived ECs1581. To test this, representative variants from the three main EHEC lineages were cloned into low copy number vector pWSK29 (pECs1581-LI; pTW14539-LI/II and p96788-LII), and assessed for their effects on T3S and motility in an *ecs1581* deletion background. All of the variants were able to induce T3S secretion and repress motility to similar levels, although the lineage II and lineage I/II variants of ECs1581 were slightly more active for these respective phenotypes (Figure 5.6A-B). These results may indicate that ECs1581 has the capacity to stimulate T3S similarly in all EHEC; although the distinct lineage divergence between these proteins may reflect real differences pertaining to the activities of these variants in their natural genetic backgrounds.

5.5.2 Congo-red binding

The ability to bind CR dye has historically been used as a marker to distinguish between pathogenic and non-pathogenic strains of bacteria (Payne and Finkelstein, 1997). More specifically, curli fibres are known to bind well to CR dye; and the expression of these surface expressed fibres have long been

associated with bacterial aggregation, adherence and biofilm formation (Barnhart and Chapman, 2006). Using CR binding as a crude phenotypic marker for curli expression, the capacity to bind CR was assessed in an *ecs1581* deletion strain *in trans* expressing wild type proteins C1493, ECED1_1787 or ECs1581; or an engineered derivative. Poor binding to CR was observed for both the *ecs1581* deletion and TUV93-0 parent (Figure 5.7). Enhanced binding was observed in those strains expressing pC1493 and pECs1581 *in trans*; but interestingly not ED1a derived ECED1_1787 (pECED1_1787) (Figure 5.7). This same pattern was observed when these strains were analysed for T3S and motility (Figures 5.5A-D). Increased binding was also observed for all the EHEC derived variants (pECs1581R20C, pECs1581K35N and pECs1581R20C+K35N) and the commensal ECED1_1787 variant engineered to harbour and RGD motif (pECED1_1787C20R) (Figure 5.7 and 5.7); a phenotype that was not observed for the empty vector control strain (pWControl). At this stage it is not known how ECs1581 up-regulates CR binding in EHEC, although an alteration in charge at the surface of the bacterial cell may help explain this phenotype. More obvious perhaps is the up-regulation of curli fibres; a well known ligand for CR and an important colonisation factor associated with virulence in both EHEC and UPEC. These preliminary observations indicate

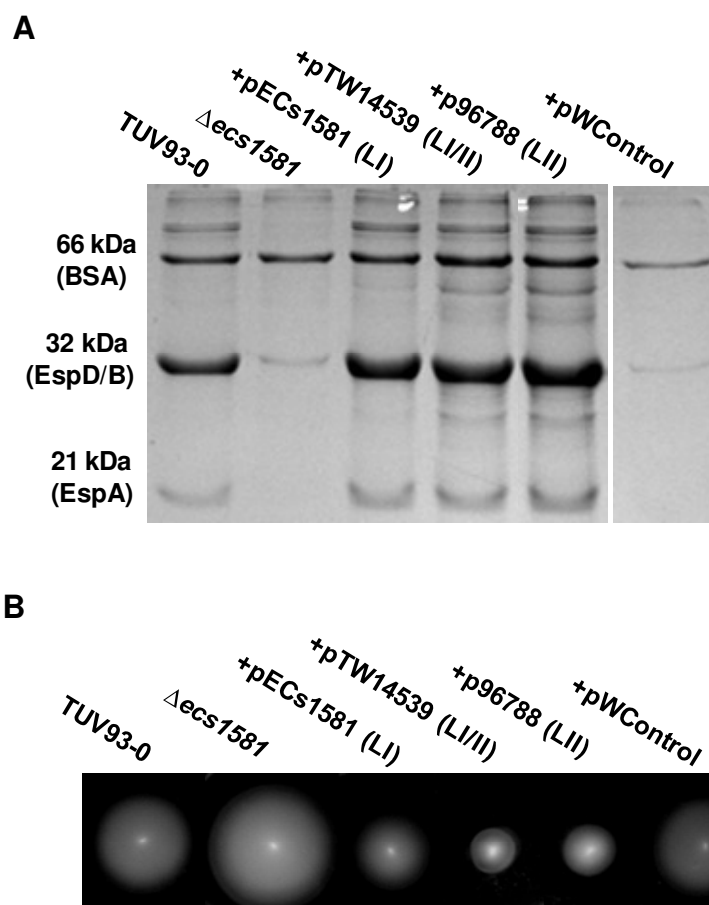


Figure 5.6. EHEC O157:H7 lineage specific variants of ECs1581 all have the capacity to stimulate T3S and repress motility. (A-B) ECs1581 was PCR amplified from strains representing the three main EHEC lineages and cloned into low copy number pWSK29 to create expression constructs pECs1581 (LI), pTW14539 (LI/II) and p96788 (LII). These variants were then analysed for their effects on T3S levels and motility in an *ecs1581* deletion background. Empty vector alone (pWControl) acted as negative control. Strains were assayed as described in Chapter 2.

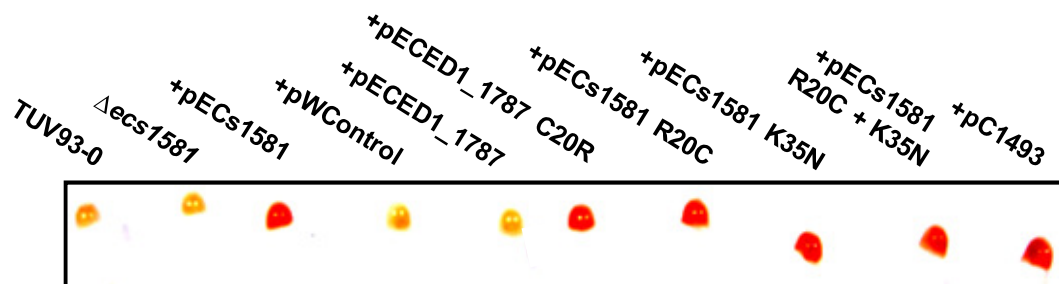


Figure 5.7. ECs1581 enhances binding to Congo-red. Natural and engineered variants of ECs1581 (annotated as shown) were cloned and expressed in an *ecs1581* deletion background and analysed for their capacity to bind Congo-red. Induced expression of ECs1581 enhanced Congo-red binding, a phenotype that was not observed for the control strain harbouring only empty vector (pWControl). Indeed, all EHEC and UPEC derived variants enhanced binding to the dye, although the commensal variant (pECED1_1787) did not appear to have this capacity. A targeted C20R mutation in ECED1_1787 (pECED1_1787C20R) was enough to then confer an increased Congo-red binding phenotype in the *ecs1581* deletion harbouring this plasmid. Strains were analysed as described in Chapter 2.

that ECs1581 is pleiotropic in that it is able to control the expression of multiple factors in EHEC including T3S, motility and surface charge.

5.6 Analysis of ECs1581 DNA binding activity

In order to investigate if ECs1581 regulation of T3S in EHEC is direct or indirect, electrophoretic mobility shift assays (EMSA) were carried out using a 291 bp *LEE1* promoter fragment (-288 to +3 with respect to the ATG translational start site of *ler*) and purified ECs1581. It was evident that the protein bound to the DNA in a concentration-dependent manner (Figure 5.8A). To examine the specificity of this interaction, ECs1581 binding to a control *gapA* promoter was also tested. Less interaction was evident at this promoter indicating that ECs1581 is a promiscuous DNA binding protein but with some degree of sequence preference (Figure 5.8B). As commensal variant ECED1_1787 was unable to activate T3S in an EHEC *ecs1581* mutant, it was a possibility that this variant may be deficient in its capacity to bind the *LEE1* promoter fragment. Purified ECED1_1787 was unable to bind the *LEE1* promoter fragment at the same concentrations as ECs1581 (Figure 5.8A). The commensal derived variant that had been engineered to harbour an RGD tri-peptide motif

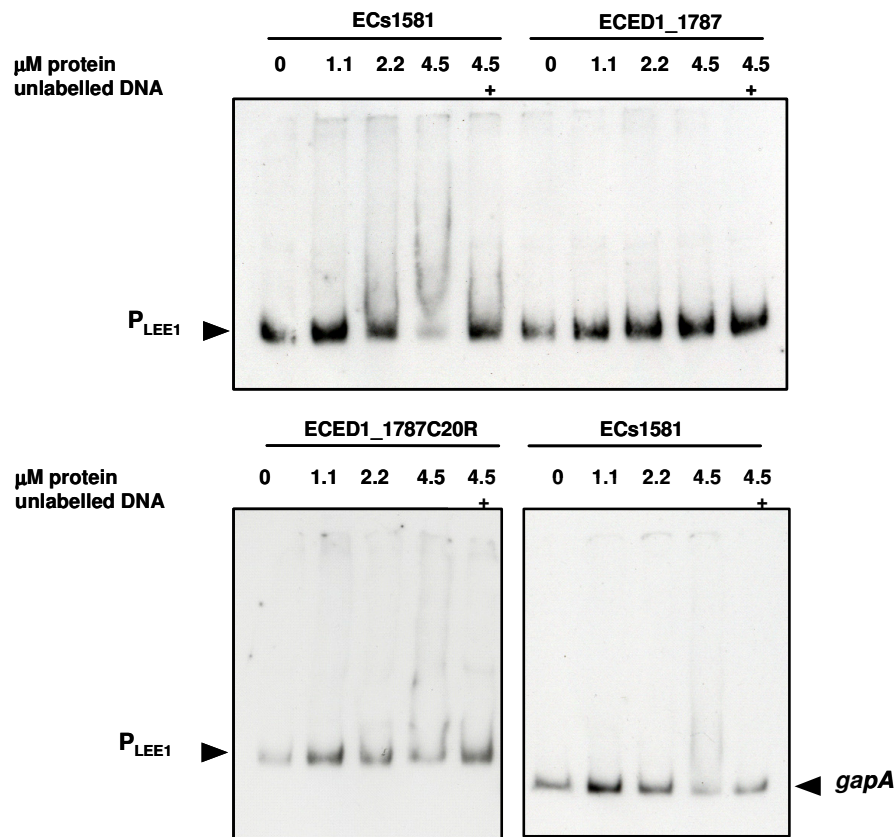


Figure 5.8. ECs1581 interacts with the *LEE1* promoter in EHEC.

(A-B) Electrophoresis mobility shift assays examining ECs1581, ECED1_1787 and ECED1_1787C20R binding to specific DNA regions. Purified proteins were added to the *LEE1* promoter at the concentrations shown. ECs1581 was also incubated with a *gapA* promoter to examine the specificity of this interaction. ECs1581, and to a lesser degree ECED1_1787C20R, but not ECED1_1787 retarded the *LEE1* promoter fragment. No shift was observed with any of the proteins in the presence of cold promoter at the highest concentration of each protein. EMSAs were kindly performed by Dr. Jai Tree as described in Chapter 2.

(ECED1_1787C20R) and which was partially active for T3S stimulation showed some evidence of binding to the same region. These results suggest that ECs1581 may be dependent on both protein-protein and protein-DNA interactions for full activity.

5.7 *In vitro* transcription assay (IVT)

In vitro transcription assays were used to determine if ECs1581 was able to activate RNA transcription from the *LEE1* promoter. However, addition of the purified protein repressed transcription *in vitro* in a concentration dependent manner (Figure 5.9A). This supports the EMSA results indicating that ECs1581 interacts with this promoter region, although the activity of the protein is repressive under these *in vitro* conditions. This activity was specific as the *gapA* transcript levels were unaffected (Figure 5.9A). As H-NS is known to be a silencer of gene expression, including from the *LEE1* promoter, the capacity of ECs1581 to activate transcription from the *LEE1* promoter was tested in the presence of H-NS. H-NS repression of the *LEE1* promoter was shown to occur

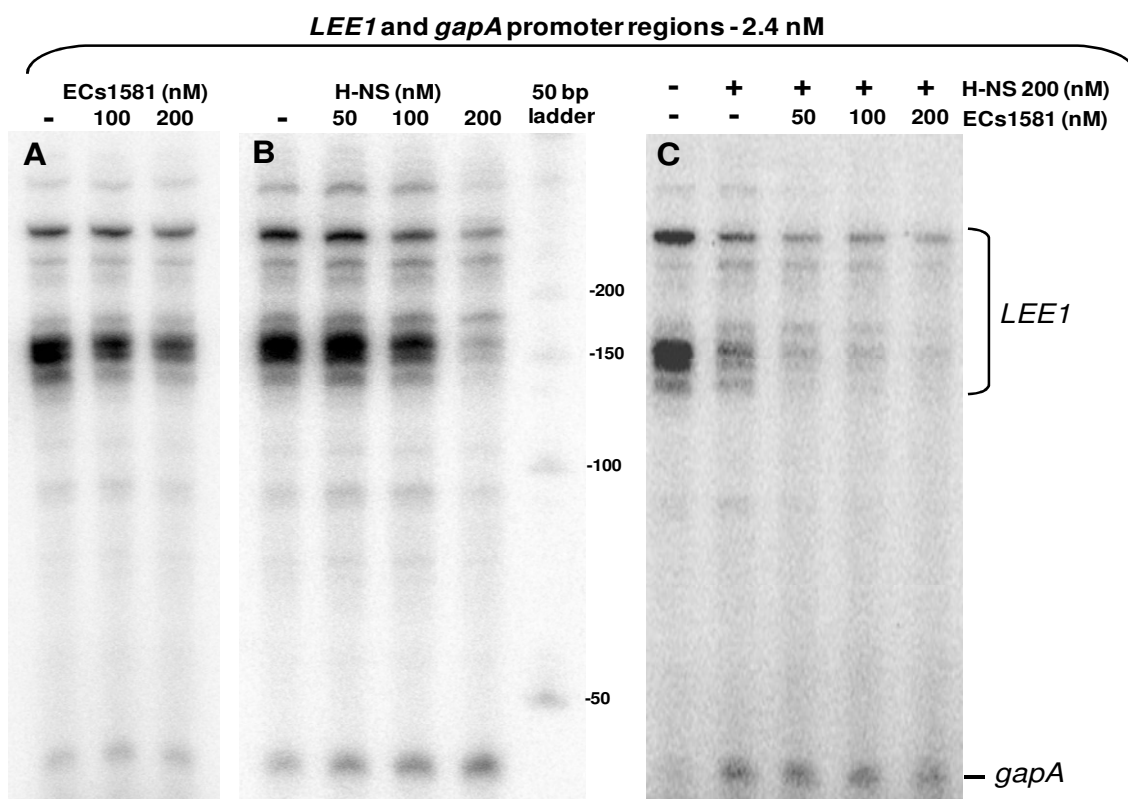


Figure 5.9 EMS1581 represses *LEE1* transcription in an *in vitro* transcription set up (A) Reactions were carried out with two promoter regions, *LEE1* and *gapA* pre-incubated with the concentrations of EMS1581 as shown. The main transcript initiated from the *gapA* promoter is estimated to be 39 bp and is indicated. The *LEE1* promoter region in EHEC has been shown to contain at least two promoters (Sperandio *et al.*, 2002); a distal promoter (P1) and proximal promoter (P2) that would give transcripts of 163 bases and 32 bases respectively in this assay. Corresponding transcripts matching those expected from P1 of (150-160 bases) and also a larger transcript (~250 bases) are observed and are labelled as *LEE1* transcripts. EMS1581 reduces the levels of the *LEE1* ascribed transcripts but not that of *gapA*. (B) H-NS was pre-incubated with the promoter regions at the concentrations shown and this repressed *LEE1*, but not *gapA* associated transcripts in a concentration-dependent manner. (C) H-NS and EMS1581 were co-incubated with the two promoter regions at the concentrations shown. EMS1581 repressed generation of *LEE1* but not *gapA* transcripts in the presence of H-NS. IVT assays were kindly carried out by Mr Ronen Rosenblum and Dr. Thamarai Scheniders at the Queen's University Belfast.

in concentration-dependent manner (Figure 5.9B) and addition of ECs1581 enhanced this repression (Figure 5.9C). Again, this ECs1581 activity was contrary to that demonstrated for *ecs1581* deletion and complementation, but supports an interaction of ECs1581 with the *LEE1* promoter under the conditions used in this assay.

5.8 Analysis of ECs1581 mechanism of action

To examine how ECs1581 activates T3S, its capacity to stimulate a *LEE1::gfp* transcriptional fusion was studied in different genetic backgrounds. Population fluorescence levels from the *LEE1* fusion were reduced significantly in an *ecs1581* mutant (Figure 5.10A). The LEE-encoded regulator (Ler) is required for normal levels of *LEE1* induction leading to T3S and so the ability of ECs1581 to stimulate *LEE1* expression in the absence of *ler* was investigated. Population fluorescence levels from the *LEE1* reporter were determined in EHEC strains ZAP198 (WT) and ZAP1004 (Δler) following induction of *ecs1581* (pECs1581) (Figure 5.8B-C). ECs1581 was able to stimulate *LEE1* expression in the absence of *ler* (Figure 5.10C). A strain was also tested (ZAP1327) in which the *LEE1* promoter had been replaced by a cassette that allows constitutive expression of *ler* but without sequences required for the complex transcriptional regulation of

this operon, including putative sequences required for *ler* auto-regulation (Mellies *et al.*, 2008; Yerushalmi *et al.*, 2008). Constitutive *ler* transcription in this replaced promoter strain was confirmed by RT-PCR (Figure 5.10E). ECs1581 induction of the *LEE1* reporter fusion in this strain background was detectable but remained at a constant level throughout the growth curve (Figure 5.10D). This was in stark contrast to the activation kinetics measured for the *LEE1* fusion in the strain with an unaltered *LEE1* promoter on the chromosome. In this wild type background, the stimulation of the *LEE1* fusion increased exponentially with optical density, where this increase was restricted when the *LEE1* chromosomal promoter was replaced effectively providing a consistent level of *ler* expression. Therefore, ECs1581 only completely activates *LEE1* in the presence of Ler in a background that allows normal Ler auto-regulation. To investigate this further, the impact of induced ECs1581 (pECs1581) and Ler (pLer) on T3S levels were examined. Induced levels of Ler or ECs1581 were able to increase T3S in an *ecs1581* mutant background while only Ler induction was able to induce T3S expression in the strain with the replaced *LEE1* promoter (Figure 5.10F). As ECs1581 does not activate T3S in this strain, it indicates as above, that the induction of T3S has to occur through the activation of the *LEE1*

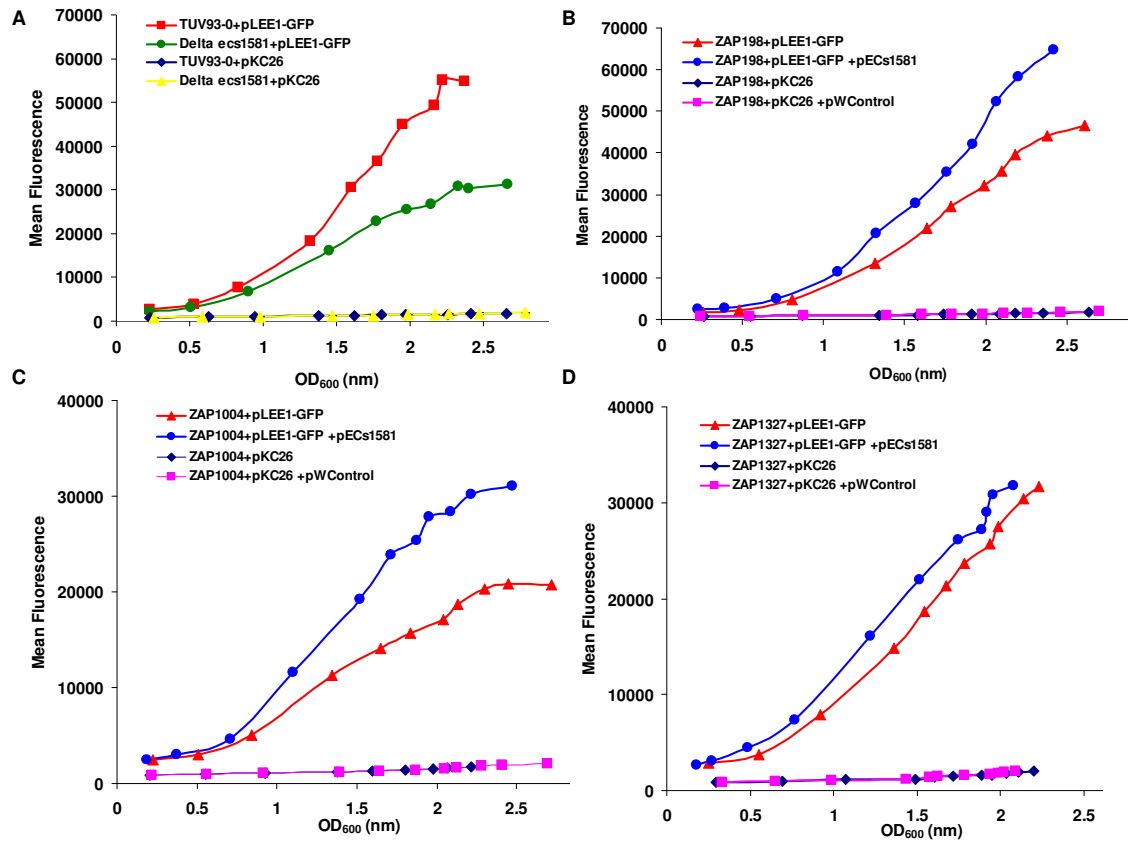


Figure 5.10 ECs1581 stimulates T3S via *LEE1* (A) Measurement of *LEE1* promoter activity in EHEC strain TUV93-0 and isogenic *ecs1581* mutant. (B) Measurement of *LEE1* promoter activity in EHEC strain ZAP198; with and without induced expression of *ecs1581* (pECs1581). (C) Measurement of *LEE1* promoter activity in a *ler* deletion strain (ZAP1004); with and without induced expression of pECs1581. (D) Measurement of *LEE1* promoter activity in an EHEC strain (ZAP1327) engineered to constitutively express *ler* on the chromosome; with and without induced expression of pECs1581. Empty vector controls in each strain background acted as a control for auto-fluorescence. All transcriptional fusion experiments were gratefully carried out by Maryia Karpiyevich as detailed in Chapter 2.

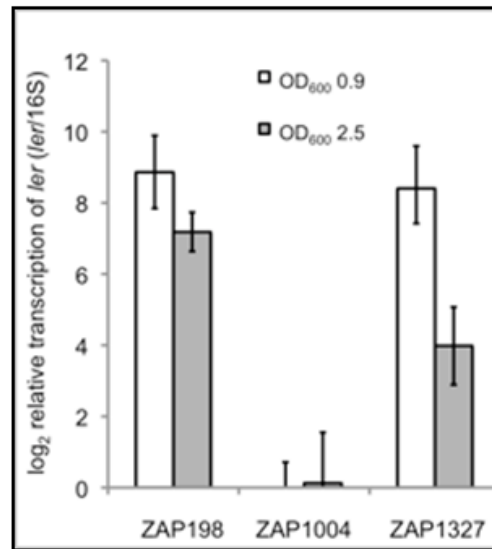
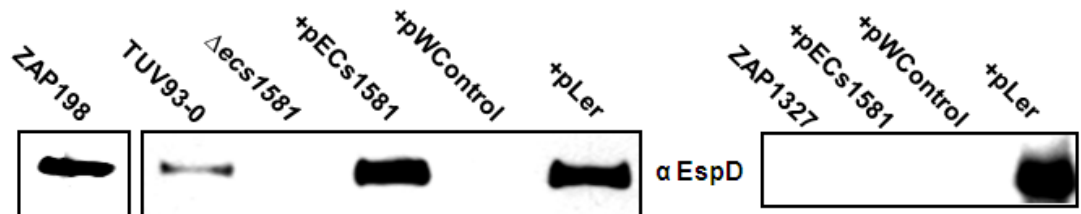
E**F**

Figure 5.10 continued (E) RT-PCR measurement of *ler* transcription levels in EHEC strains ZAP198 (wild type), ZAP1004 (Δ ler) and ZAP1327 (Δ LEE1 promoter). Of note, *ler* transcription is comparable in the wild type and replaced *LEE1* promoter strain during exponential growth (OD₆₀₀ 0.9), however, a small dip in expression can then be observed in the constitutively expressed *ler*/*LEE1* strain as the cells enter into stationary phase (OD₆₀₀ 2.5). As ongoing research in the laboratory is identifying multiple sRNAs that can act on *LEE* transcripts, including *LEE1*, this data may indicate that a change is occurring in the post-transcriptional control of *ler* at the transition from exponential phase to stationary phase growth; coinciding with the OD at which the dip in expression occurs. **(F)** Western blot detection of EspD in an EHEC *ecs1581* deletion and replaced *LEE1* promoter mutant (ZAP1327) *in trans* expressing Ler (pLer) or ECs1581 (pECs1581). Empty vector alone (pWControl) was used as a negative control. The above RT-PCR experiment was kindly performed by Maryia Karpiyevich as described in Chapter 2.

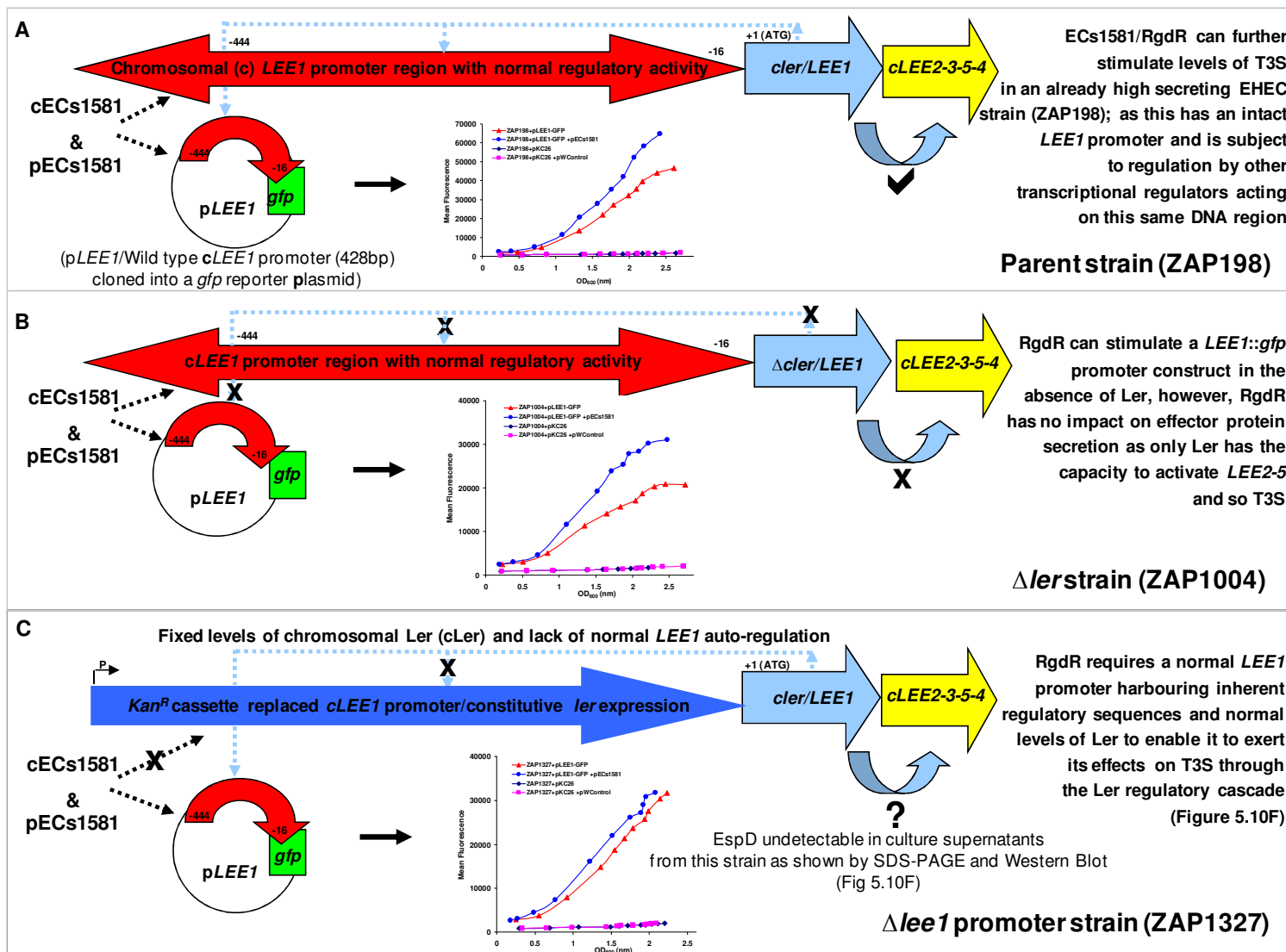


Figure 5.11 (A-C) Schematic summary of how ECs1581 acts to positively regulate *LEE1* and T3S in EHEC (a full description is given overleaf/page 175)

Figure 5.11 (A-C) Schematic interpretation of how ECs1581/RgdR acts to positively regulate *LEE1* and T3S in EHEC. To examine how ECs1581 activates T3S, its capacity to stimulate a 428 bp (-16 to -444 in relation to *ler* ATG translational start codon) *LEE1::gfp* transcriptional fusion was studied in different genetic backgrounds (Figures 5.10A-D/5.11A-C). The LEE-encoded regulator (Ler) is required for normal levels of *LEE1* induction leading to T3S and so the ability of ECs1581 to stimulate *LEE1* expression in the absence of *ler* was investigated (Figure 5.10C/5.11B). Population fluorescence levels from the *LEE1* reporter were determined in EHEC strains ZAP198 (WT) and ZAP1004 (Δ *ler*) following induction of *ecs1581* (pECs1581) (Figure 5.10B-C/5.11A-B). ECs1581 was able to stimulate *LEE1* expression in the absence of *ler* (Figure 5.10C/5.11B). A strain was also tested (ZAP1327) in which the *LEE1* promoter had been replaced by a cassette that allows constitutive expression of *ler* but without sequences required for the complex transcriptional regulation of this operon, including putative sequences required for *ler* auto-regulation (Figure 5.10D/5.11C) (Mellies *et al.*, 2008; Yerushalmi *et al.*, 2008). Constitutive *ler* transcription in this replaced promoter strain was confirmed by RT-PCR (Figure 5.10E). ECs1581 induction of the *LEE1* reporter fusion in this strain background was detectable but remained at a constant level throughout the growth curve (Figure 5.10D and 5.11C). This was in stark contrast to the activation kinetics measured for the *LEE1* fusion in the strain with an unaltered *LEE1* promoter on the chromosome (Figure 5.10B/5.11A). In this wild type background, the stimulation of the *LEE1* fusion increased exponentially with optical density, where this increase was restricted when the *LEE1* chromosomal promoter was replaced effectively providing a consistent level of *ler* expression. Therefore, ECs1581 only completely activates *LEE1* in the presence of Ler in a background that allows normal Ler auto-regulation. To investigate this further, the impact of induced ECs1581 (pECs1581) and Ler (pLer) on T3S levels were examined. Induced levels of Ler or ECs1581 were able to increase T3S in an *ecs1581* mutant background while only Ler induction was able to induce T3S expression in the strain with the replaced *LEE1* promoter (Figure 5.10F). As ECs1581 does not activate T3S in this strain, it indicates as above, that the induction of T3S has to occur through the activation of the *LEE1* (Flockhart *et al.*, 2012). A blue dotted line indicates probable interactions between Ler and its target DNA. A blue dotted line with a black cross indicates that this regulation is in some way compromised. A black dotted line indicates probable interactions between chromosomal (c) derived and plasmid (p) derived ECs1581 and its target DNA. A black dotted line with a black cross indicates that this regulation is in some way compromised. Remaining symbols, including directional arrows, are annotated as shown.

promoter and the Ler auto-stimulatory circuit to promote expression of the remaining *LEE2-5* operons and T3S effector protein secretion. A schematic model/interpretation of ECs1581's considered activity is shown in Figure 5.10G (Flockhart *et al.*, 2012).

5.9 Discussion

Through a combination of deletion and complementation analyses, *ecs1581* was identified as a gene able to up-regulate T3S in EHEC (chapter 4). ECs1581 encodes a putative 99 amino acid protein of 11.7 kDa with two main variants present in EHEC O157:H7 strains (Figure 5.2A), although multiple minor variants are also identified. This protein has little homology with established bacterial regulators although there is a region of homology with Ler over a region considered to link the oligomerisation and DNA binding regions of this global regulator (Figure 5.1) (Mellies *et al.*, 2008).

Three lineages of EHEC O157:H7 have been proposed based on genome sequences (Kim *et al.*, 1999; Zhang *et al.*, 2007). In the present study, two main sequence types of ECs1581 were identified, LI and LII & LI/II, the latter having identical protein coding sequences. EHEC O157:H7 lineage specific variants of

ECs1581 (LI) from strains TW14539 (LI/II) and 96788 (LII) were cloned and assessed for their effects on T3S and motility. All the variants were able to induce T3S and repress motility although slight variation in induction levels was observed (Figure 5.6A-B). Sequences immediately upstream of these two variants also differ (data not shown) but it is not known if expression levels of these two variants differ as these were not assessed. Variation in activity and/or expression of ECs1581 could therefore contribute to differences in T3S expression demonstrated between EHEC O157:H7 strains (Rashid *et al.*, 2006; Roe *et al.*, 2003; Son *et al.*, 2002; Yang *et al.*, 2009). Of note, the genetic variability among the different EHEC O157:H7 lineages is also reported to be associated with regions adjacent to *pchD* and *pchE* genes on OI-43/48 (the tellurite resistance and adhesin-conferring islands) and OI-51 respectively (Zhang *et al.*, 2007 and Yang *et al.*, 2009).

Other variants of the protein from different *E. coli* backgrounds exhibited markedly different capacities to regulate T3S. For example the variant from the UPEC *E. coli* strain CFT073 (C1493) induced T3S to higher levels than any other variant tested, while a variant (ECED1_1787) from the B2 commensal *E. coli* strain ED1a, a close relative of CFT073, had no discernible effect on T3S; despite

varying at only 6 residues compared to ECs1581. This natural variant therefore provided insights into essential residues for ECs1581 function. One of the natural changes in ECED1_1787 was an R20C (CGD) in the EHEC conserved RGD motif, and studies have demonstrated that this motif can play an important role in mediating cell adhesion and ligand binding (Ruoslahti, 1996). Substitution of this residue to generate an RGD motif in ECED1_1787 (C20RGD) produced a protein that now had the capacity to activate T3S. Based on this finding, ECs1581 was termed RgdR. As ECED1_1787 was the only variant with an N35, in comparison to K35 in all EHEC ECs1581 variants and CFT073 C1493, an ECs1581 K35N variant was constructed and this also had reduced potential to activate T3S. Combining a double R20C+K35N change in ECs1581 resulted in a variant that had the lowest capacity of any variant tested to stimulate T3S (Figure 5.5C). While some of these changes will impact directly on activity, it is appreciated that the small size of RgdR means that some changes will also affect protein conformation and stability.

As bacterial motility was also altered by deletion and complementation of *ecs1581*, the effects of the different variants on motility were examined. It was interesting to note that regardless of the effect on T3S, all variants in RgdR had a

marked impact on motility. In general there was an inverse correlation between T3S and motility which was also apparent for several of the engineered variants. However, there were clear exceptions to this, for example ECs1581 R20C+K35N had a weak capacity to activate T3S yet unusually, showed the strongest capacity to inhibit motility. Additional work confirmed that this regulation of T3S and motility was independent of *grlA*, a LEE-encoded regulator previously demonstrated to repress motility on induction of T3S (Iyoda *et al.*, 2006). RgdR is therefore identified as a novel regulator able to coordinate T3 and motility expression. Also of note was the finding that expression of different variants could increase bacterial binding to Congo red, indicating that RgdR is likely to control the expression of multiple surface components. Although the mechanism of this Congo red binding was not investigated, recent work by Lee *et al* (2011) has demonstrated that EDL933 is able to produce curli fibres. Taken together, the pleiotropic phenotypes displayed by RgdR may explain the presence of variants of this regulatory protein in strains without a LEE-encoded T3S system.

RgdR was able to stimulate *LEE1* expression independently of Ler and required a normal *LEE1* promoter region for this activation. A construct that resulted in

constitutive expression of *LEE1* could be induced to secrete EspD by increased expression of Ler *in trans* but not by induced expression of RgdR *in trans*. In this background, RgdR was limited in its capacity to induce a *LEE1::gfp* reporter fusion construct. An explanation for this is that even though RgdR can activate the *LEE1* promoter in the absence of Ler, Ler expression and the Ler auto-induction cycle is required for the activation of T3S as Ler, and not RgdR, activates transcription of *LEE2-5* encoding the T3S system and secreted proteins. One explanation for the data is that ECs1581 could inhibit Ler auto-repression at the *LEE1* promoter. It has been proposed that high Ler levels allow binding of Ler to a low affinity site which may repress transcriptional activation. The kinetics of *LEE1* activation in the presence of induced ECs1581 would fit this model. Repeated attempts to purify functional Ler to test possible interactions were unsuccessful. Of note, ECs1581 does share homology with Ler over a linker region (Figure 5.1) that could be involved in this interaction.

RgdR also bound directly to the *LEE1* promoter region and may require this function for its activity, supported by the finding that a non-binding variant did not stimulate T3S. Purified RgdR did not activate *LEE1* transcription *in vitro*, in fact it repressed transcription. Moreover, there was no evidence that RgdR

could relieve H-NS repression *in vitro*, as suggested for Ler (Bustamante *et al.*, 2001; Stoebe *et al.*, 2008); again ECs1581 enhanced H-NS repression of transcription from the *LEE1* promoter. It is important to note that *in vitro* transcriptional assays do not involve the multiple activators and repressors demonstrated to act on this promoter (Kendall *et al.*, 2010; Tree *et al.*, 2009) and so it is highly plausible that RgdR may act as a co-factor with one or more of these to activate transcription. One regulatory network investigated was the quorum sensing two-component system, *qseBC*, as this controls both T3S and motility (Hughes *et al.*, 2009; Kostakioti *et al.*, 2009). However, RgdR was able to significantly increase T3S levels in a TUV93-0 *qseC* deletion mutant (data not shown). Another possibility that remains to be tested is that RgdR works with the Pch family of LEE activators (Iyoda *et al.*, 2004; Abe *et al.*, 2008; Yang *et al.*, 2009), although there is no significant homology between the two, in particular to the region truncated in the smaller Pch variants (PchC-E) (Abe *et al.*, 2008).

With the exception of *pchD*, all *pch* alleles are encoded on prophage elements, and it is hypothesised that multiple prophages in the chromosome of EHEC O157:H7 may be able to recombine to affect the total copy number of *pch* genes; thus potentially impacting on levels of LEE gene expression (Iyoda and

Watanabe, 2004; Ohnishi *et al.*, 2001). As with the Pch activators of LEE, all members of the RgdR family of proteins are prophage associated. For example, EHEC strain EDL933 harbours two RgdR truncates (both 63 amino acids), *z1197* and *z1636* that are located on OI-43 and OI-48 respectively (Table 5.1). In EHEC strain Sakai only one truncate exists (63 amino acids), *ecs5415*, that is located on Sakai prophage-like element 1 (SpLE1). In EHEC O26:H11 strain 11368, four copies of RgdR can be found (Table 5.1); three full length (99 amino acids) variants and one truncate (63 amino acids) which is located on OI-48. Perhaps, as proposed for the Pch proteins, genome integration of CP-933C (OI-51) and similar prophages may provide a mechanism whereby RgdR proteins can additively influence levels of LEE gene expression in EHEC; however, further experimental investigation would be required to address this.

OI-51 was demonstrated to be important for EHEC O157:H7 colonisation and persistence in sheep. This could be a direct result of RgdR acting on T3S regulation, but may also reflect coordinated regulation with other genes on this island. Future work should aim to look at the context of RgdR along with the function of other genes surrounding this genetic region. In summary, this research has defined an important region (OI-51) necessary for EHEC O157:H7

colonisation and persistence in sheep, and has identified a completely new family of small bacterial regulators that have the capacity to control surface factor expression in *E. coli*.

Table 5.1 List of ECs1581 orthologous and paralogous proteins present in *E. coli* and *Shigella spp* strains

ECs1581 variant (pathoty pe/strain/locus tag)	Length (aa)	Residues found at positions 19-22	% sequence similarity with ECs1581/ overlap region (aa)	Comments/references
EHEC strains Sakai/EDL933				
ECs1581	99	MRGD	100 over 99	These prototypical EHEC strains were isolated during large outbreaks in Japan and the US from radish sprouts and undercooked hamburger meat respectively. Perna <i>et al.</i> , 2001; Hayashi <i>et al.</i> , 2001
ECs5415 (SpLE1)	63	GALV	39 over 41	
Z1197 (OI-43)	63	GALV	39 over 41	
Z1636 (OI-48)	63	GALV	39 over 41	
Commensal <i>E. coli</i> strain ED1a (O81)				
ECED1_1787	99	MCGD	94 over 99	Commensal <i>E. coli</i> strain isolated from the faeces of a healthy individual. Clermont <i>et al.</i> , 2008
ECED1_2647	99	IPGA	31 over 99	
ECED1_3312	97	MRGD	67 over 97	
ECED1_3445	63	GALA	42 over 45	
Atypical EPEC strain E110019				
EcE110019_0327	99	MRGD	92 over 99	Atypical EPEC strain E110010 was isolated from an outbreak in a school in Finland. This strain lacks the EAF plasmid. Rasko <i>et al.</i> , 2008, Craig. J. Venter Institute
EcE110019_4878	99	TRGD	80 over 99	

Table 5.1 cont

***Shigella flexneri* 5 strain 8401**

SFV_1149	99	TRGD	88 over 99	Nie <i>et al.</i> , 2006
----------	----	------	------------	--------------------------

***Shigella flexneri* strain 2002017**

SF _{xv} _1290	99	TRGD	88 over 99	<i>S. flexneri</i> 2002017 is an emerging multi drug resistant serotype X variant of <i>Shigella</i> , isolated in China from an epidemic clone. Ye <i>et al.</i> , 2010
------------------------	----	------	------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------

***Shigella dysenteriae* strain 1012**

Sd1012_1853	99	MNGD	85 over 99	<i>S. dysenteriae</i> 1012 was isolated in Bangladesh and is representative of type 4 <i>S. dysenteriae</i> that are becoming more prevalent in humans infections. Studies have shown this to be one of the most virulent <i>S. dysenteriae</i> strains identified Rasko <i>et al.</i> , 2008, J. Craig.Venter Institute
Sd1012_2233	99	MRGD	79 over 99	

***E. coli* sp 3_2 53FAA**

ESAG_00330	99	MNGD	84 over 99	<i>E. coli</i> sp. 3_2_53FAA was isolated from an intestinal biopsy specimen taken from a 52-year old male with active Crohn's disease.
------------	----	------	------------	-----------------------------------------------------------------------------------------------------------------------------------------

Table 5.1 cont**NMEC *E. coli* strain IHE3034**

ECOK1_0888	99	MNGD	84 over 99	Recently sequenced newborn meningitis <i>E. coli</i> strain. Moriel <i>et al.</i> , 2010
------------	----	------	------------	------------------------------------------------------------------------------------------

UPEC strain CFT073

C1493	99	MRGD	83 over 99	Prototypical UPEC <i>E. coli</i> strain CFT073. Mobley <i>et al.</i> , 1990
C4520	63	GALA	42 over 41	

UPEC strain IAI39

ECIAI39_2021	99	MNGD	83 over 99	Touchon <i>et al.</i> , 2009
ECIAI39_0509	95	SLGD	44 over 92	

EHEC O26:H11 strain 11368

ECO26_1651	99	TRGD	81 over 99	Ogura <i>et al.</i> , 2009
ECO26_3111	99	TRGD	78 over 99	
ECO26_2322	99	IPGA	29 over 99	
ECO26_1322	63	GALV	39 over 41	

EHEC strain TW14539

ECSP_1496	99	TRGD	81 over 99	A spinach outbreak isolate of EHEC O157:H7 that is considered to be hyper-virulent. Abu-Ali <i>et al.</i> , 2010
ECSP_1303	63	GALV	39 over 41	

EHEC O103:H2 strain 12009

ECO103_1241	99	TRGD	81 over 99	Ogura <i>et al.</i> , 2009
-------------	----	------	------------	----------------------------

Table 5.1 cont**EPEC O55:H7 CB9615**

G2583_1395	99	TRGD	81 over 99	This strain is considered to be the closest ancestor of EHEC O157:H7 clones. Zhou <i>et al.</i> , 2010
------------	----	------	------------	--------------------------------------------------------------------------------------------------------

EAEC O42

EC042_1760	99	TRGD	80 over 99	Prototypical enteroaggregative <i>E. coli</i> strain O42. Chaudhuri <i>et al.</i> , 2010
EC042_2163	99	RPGD	68 over 99	

EHEC O111:H- strain 11128

ECO111_2919	99	TRGD	78 over 99	Ogura <i>et al.</i> , 2009
-------------	----	------	------------	----------------------------

Chapter 6

General discussion

6.1 General discussion

6.1.1 *E. coli* O157:H7: current understandings in relation to this research

EHEC is an important zoonotic pathogen harbouring prophage-encoding Shiga toxins that are associated with severe GI and systemic disease in humans (Karmali, 2004; Nataro and Kaper, 1998). Ruminants are considered to be the most important reservoirs for EHEC, particularly cattle and sheep which shed the organism in their faeces (La Ragione *et al.*, 2009). The predominant pathogenic serotype in North America, parts of Asia and the United Kingdom is O157:H7 (Armstrong *et al.*, 1996; Besser *et al.*, 1999; Caprioli *et al.*, 2005). The horizontally acquired LEE PAI encodes a type III secretion (T3S) system, an essential virulence factor for EHEC which allows the direct injection of bacterial effector proteins into host cells to subvert host cell signalling pathways and promote bacterial attachment (Hemrajani *et al.*, 2010; Kenny *et al.*, 1997; Newton *et al.*, 2010). Whilst T3S has been shown to be critical for EHEC O157:H7 colonisation in the ruminant host (Naylor *et al.*, 2005), levels of T3S can vary significantly amongst different EHEC isolates (Rashid *et al.*, 2006; Roe *et al.*, 2003; Son *et al.*, 2002; Yang *et al.*, 2009). Such variation in T3S regulation may have important implications for bacterial persistence and shedding in the bovine host and for human infection (Chase-Topping *et al.*, 2008).

Studies comparing EHEC O157:H7 genome sequences have identified three lineages of the organism (Kim *et al.*, 1999; Zhang *et al.*, 2007). While all lineages can be isolated from cattle, lineage I (LI) and lineage I/II (LI/II) strains are more frequently associated with human clinical disease than lineage II (LII) strains. Further, EHEC O157:H7 strains associated with outbreaks of severe illness in the United States of America have been shown to belong to a single subtype of LI/II, designated SNP clade 8 (Laing *et al.*, 2009; Manning *et al.*, 2008). Several studies have attempted to understand the genetic basis for the apparent differences in the epidemiology and virulence among EHEC O157:H7 genotypes and recent studies have shown that LI and LI/II strains produce higher levels of Shiga toxins than bovine-associated LII strains (Zhang *et al.*, 2010). Similarly, hyper-virulent SNP clade 8 strains have been shown to have higher levels of LEE expression than EHEC O157:H7 strain Sakai from SNP clade 2 (Abu-Ali *et al.*, 2010).

A large number of factors have been shown to influence expression of LEE in both EPEC and EHEC (Kendall *et al.*, 2010; Tree *et al.*, 2009). In both pathotypes, the LEE-encoded regulator (Ler) is expressed from the first LEE operon (*LEE1*) and is required to activate expression of all five main LEE operons (*LEE1-5*)

(Elliott *et al.*, 2000; Sharma and Zuerner, 2004). However, at higher concentrations, Ler is considered to be repressive at *LEE1* (Berdichevsky *et al.*, 2005; Yerushalmi *et al.*, 2008). To date, most of the genetic and environmental control of T3S has been shown to occur through *LEE1* and Ler activation, including an activator and repressor combination within the LEE, GrlA and GrlR, that are involved in the reciprocal regulation of LEE and flagella expression (Barba *et al.*, 2005; Huang and Syu, 2008; Islam *et al.*, 2011; Iyoda *et al.*, 2006; Jimenez *et al.*, 2010). In EPEC, LEE expression is thought to be positively regulated by the plasmid-encoded regulator C (PerC) locus and a series of PerC homologue (Pch) proteins; encoded by *pchA-E* and *pchX* in EHEC O157:H7. These genes are usually associated with cryptic prophage regions encoding effector proteins that are exported by the T3S system and the Pch regulators co-ordinate the expression of these horizontally-acquired effectors with the LEE-encoded T3S system (Abe *et al.*, 2008; Iyoda and Watanabe, 2004; Porter *et al.*, 2005; Yang *et al.*, 2009). Indeed, the integration and recombination of horizontally acquired elements in EHEC O157:H7 have not only facilitated phage evolution, but have driven the acquisition of virulence attributes in these bacteria and the complex regulatory networks that control their expression. These dynamic activities of bacteriophage have sparked new interest in light of

the discovery that many sequenced bacterial genomes harbour multiple prophages carrying a wide range of genes; in particular genes related to bacterial colonisation and host pathogenesis. However, as many prophages often contain various genetic defects, they have previously been regarded as mere genetic vestiges (Asadulghani *et al.*, 2009). Recent studies, including this one, have attempted to address the functionality of these prophages in EHEC O157:H7; and it has been reported for O157:H7 strain Sakai that many of the apparently defective prophages are in fact inducible and can be released from O157:H7 cells as particulate DNA, with some even being transferred to other *E. coli* strains (Asadulghani *et al.*, 2009). Moreover, the genotypic diversity amongst different strains of EHEC has been attributed, by in large, to phage-related DNA fragments (Ohnishi and Hayashi, 2002; Zhang *et al.*, 2007); with these horizontally acquired elements being termed 'O-islands' (OIs) in the sequenced prototypical EHEC O157:H7 strain, EDL933 (Perna *et al.*, 2001).

6.1.1.1 Summary of research findings

A primary goal of this PhD research was to phenotypically characterise prophage associated regions (OIs) in EHEC O157:H7; to try and uncover novel horizontally acquired regulators that confer on the bacterium determinants that

help it to survive and colonise in its mammalian host. An OI-51 deletion mutant was found to have lower levels of LEE expression *in vitro* and subsequent testing demonstrated that this genomic island contributes to ruminant colonisation and persistence. A novel regulator, termed RgdR, was identified on OI-51 and shown to control both LEE expression and motility; independently of the global LEE and motility regulator, GrlA, and the major QS regulator, QseC that regulates flagella gene expression and pedestal formation in EHEC (Barba *et al.*, 2005; Clark and Sperandio, 2005; Iyoda *et al.*, 2006; Kostakioti *et al.*, 2009; Sperandio *et al.*, 2002; Jobichen *et al.*, 2007). *In vitro* transcription assays could not substantiate RgdR as a classical transcriptional activator, where transcription from the *LEE1* promoter was not evident in the presence of RNAP; in fact RgdR appeared to slightly repress transcription; although this could have been non-specific given how sensitive IVT systems can be to slight changes in reaction conditions. Additionally, RgdR was also unable to stimulate *LEE1* expression in the presence of H-NS, implying that it does not function to activate LEE gene expression by de-repressing H-NS as proposed for Ler (Bustamante *et al.*, 2001; Stoebe *et al.*, 2008); or requires the activity of other protein(s) to do so. Again, as IVT systems do not take into account the presence of other proteins and transcription factors that act on the LEE regulatory

promoter region, RgdR likely activates *LEE1* transcription in the presence of these other factors. *LEE1::gfp* reporter constructs demonstrated that RgdR has the capacity to induce *LEE1* expression in the absence of the master LEE regulator, *ler*, but requires a normal *LEE1* promoter region for this activation. A construct that resulted in constitutive expression of *LEE1* could be induced to secrete EspD by increased expression of *Ler* *in trans* but not by induced expression of RgdR *in trans*. In this background, RgdR was limited in its capacity to induce a *LEE1::gfp* reporter fusion construct. One explanation for the data is that RgdR could inhibit *Ler* auto-repression at the *LEE1* promoter. It has been proposed that high *Ler* levels allow binding of *Ler* to a low affinity site which may repress transcriptional activation (Mellies *et al.*, 2008; Yerushalmi *et al.*, 2008). The kinetics of *LEE1* activation in the presence of induced RgdR would fit this model. Repeated attempts to purify functional *Ler* to test possible interactions were unsuccessful, although RgdR does share slight homology with *Ler* at a functionally important region spanning *Ler*'s oligomerisation linker domain that could be involved in this interaction (Mellies *et al.*, 2008; Yerushalmi *et al.*, 2008). RgdR did demonstrate a low affinity and non-specific interaction with DNA by EMSA, a property that was absent from a commensal variant that did not stimulate T3S; and so it is possible that RgdR acts as a co-

factor with one or more regulators of T3S to control transcription from the *LEE1* promoter.

6.1.1.2 OI-51 encoded RgdR proteins are present in diverse phylogenetic groups of *E. coli*

RgdR proteins have the capacity to influence pleiotropic phenotypes (including T3S, motility and cell surface charge) and so may support a more global role in the cell and may explain the presence of these regulatory proteins in strains without a LEE-encoded T3S system. Indeed, a cloned UPEC variant of RgdR was able to stimulate T3S in an EHEC background to levels higher than any other variant, including other EHEC O157:H7 derived variants; despite UPEC lineages not harbouring a classical LEE-encoded T3SS. This may indicate that in specific genetic backgrounds, for example in EHEC strains and *Shigella* strains, RgdR proteins have been adapted to control T3S. LEE-encoded T3SSs and flagella systems are evolutionary and structurally related and so this is conceivable (Blocker *et al.*, 2003; Erhardt *et al.*, 2010). Intriguingly, RgdR proteins are also found in commensal *E. coli* strains, begging the question as to the function(s) of these small proteins in these genetic backgrounds.

Comparative genome analyses have revealed that bacterial recombination, including the acquisition of horizontally-acquired prophage elements, is largely responsible for the variation in gene content between different strains of *E. coli* (Kim *et al.*, 1999; Tenaillon *et al.*, 2010; Touchon *et al.*, 2009; Zhang *et al.*, 2007). Indeed, the average *E. coli* genome is estimated to contain over 4000 genes, however, only 2000 of these genes are conserved amongst all strains; indicating a stark degree of plasticity in these strains at the genome level (Hendrickson, 2009; Tenaillon *et al.*, 2010). Further, studies based on MLST and MLEE (multi locus enzyme electrophoresis) have demonstrated that *E. coli* clones have broad geographical and host distributions that can be illustrated by the prevalence of the four main phylogenetic groups of *E. coli* in human and animal populations (Duriez *et al.*, 2001; Escobar-Paramo *et al.*, 2004; Escobar-Paramo *et al.*, 2006; Gordon *et al.*, 2008; Ochman *et al.*, 1984; Tenaillon *et al.*, 2010). Although this variation in the prevalence of phylogenetic groups among different hosts is not attributable to the existence of host-specific strains, certain strains are considered to have strict host specificity; including and an avirulent B2 clone of the O81 type, ED1a, which is reported to be found only in humans (Clermont *et al.*, 2008; Tenaillon *et al.*, 2010). This ED1a clone was of special interest to this research as this strain harbours a number of RgdR protein variants in its

chromosome, both full length and truncated, that are also present on prophage elements in diverse pathogenic *E. coli* strains; including EHEC, atypical EPEC, UPEC, NMEC and *Shigella* strains.

The presence of OI-51 and RgdR like proteins in these *E. coli* may illustrate a diversity of adaptive paths in the different host species; for which some clusters of genes or genomic islands should be found only in a subset of strains that are favoured in specific environments; and where several alternative combinations of genes could potentially promote similar adaptations to a given environment (Hacker and Kaper, 2000; Tenaillon *et al.*, 2010; Touchon *et al.*, 2009). With this, epidemiological data and experimental studies in animal models have identified and extensively studied genes that are associated with 'virulence', and there is now growing evidence that these virulence associated genes have evolved and been maintained by selection for other roles that they have in the ecology of the bacteria; especially in commensalism (LeGall *et al.*, 2007; Levin, 1996; Tenaillon *et al.*, 2010; Touchon *et al.*, 2009). The presence of RgdR proteins in nearly all of the major phylogenetic *E. coli* groups may indicate that RgdR proteins too have important other functions (in addition to their activities on T3S, motility and surface charge) in the cell and have been maintained in these species as a result

as part of a larger prophage unit; for example on OI-51 and OI-51-like prophage elements. Phenotypic characterisation of a cloned RgdR variant from commensal ED1a demonstrated that this protein, somewhat remarkably, was unable to stimulate T3S; despite sharing the highest homology (94%) with EHEC derived RgdR. Mutation at position 20 to create an RGD motif (C20RGD) in this ED1a wild type variant (as is found in EHEC and UPEC variants of RgdR) was enough to now confer on this protein T3S stimulatory activity; and this variant was now repressive for motility. Although purely speculative, perhaps this commensal variant has selectively diverged at this critical residue and others (whilst retaining extremely high overall sequence homology to ECs1581); to 'temper/adapt' the activity of this protein in response to the specific selective pressures that this commensal strain would inevitably encounter in its human specific host (Clermont *et al.*, 2008). Of note, the ED1a wild type protein was the only variant observed to enhance motility when induced in an EHEC genetic background; all of the other variants that were tested acted to repress motility. ED1a has also previously been studied for its use as a potential probiotic (Clermont *et al.*, 2008), but rather interestingly, this commensal harbours a number of other EHEC prophage OIs in addition to OI-51 (Clermont *et al.*, 2008; Perna *et al.*, 2001; Touchon *et al.*, 2009).

Future work could look at cytokine responses in eukaryotic cells infected with a panel of wild type parent and isogenic *rgdR* deletions representing the different pathotypes of *E. coli*; and in the same strains expressing intra and inter-pathotype RgdR variants. ED1a ECED1_1787 and its engineered variant ECED1_1787C20R could prove valuable in this experimental set up; if one of these variants were able to up-regulate pro-inflammatory cytokine responses on cells and the other not, then one could possibly speculate that this commensal variant has adventitiously diverged at residue 20. Amino acid divergence at residue 20 (e.g. R20/C20/N20/T20) is apparent on comparing RgdR sequences from EHEC, UPEC, atypical EPEC, commensal ED1a and *Shigella spp* strains. The divergence between EHEC RgdR and ED1a RgdR may therefore reflect real differences in the functionality of these proteins with respect to each strains particular niche and life style. In this context, RgdR proteins cloned from *Shigella spp* would be of particular interest to test, as *Shigella* are invasive on cells, non-motile and harbour a T3SS (Ashida *et al.*, 2011; Jin *et al.*, 2002; Schroeder *et al.*, 2008; Wei *et al.*, 2003).

6.1.1.3 Future work to address the functionality of OI-51-like prophage in *E. coli*

Microarray studies in an extra-intestinal *E. coli* (e.g UPEC or NMEC) deletion strain background would make a good starting point to address the potential regulons of OI-51/RgdR in different pathotypes of *E. coli*. For UPEC specifically, an ascending UTI model is already well established in mice (Hagber *et al.*, 1983; Mobley *et al.*, 1990) and so this could be used to compare a UPEC *c1493* deletion and parent strain for *in vivo* colonisation. For EHEC, future work could characterise global gene expression patterns in a wild type and isogenic *rgdR* deletion grown under a variety of environmental conditions, for example under T3S permissive and T3S repressive growth conditions; as this should provide valuable information pertaining to possible other targets for RgdR aside from the LEE PAI. It has been hypothesised that OI-51 interacts with other prophage in the chromosome of EHEC to facilitate its own induction and replication (Asadulghani *et al.*, 2009). As RgdR controls multiple phenotypes and is encoded on OI-51, it will be interesting to see if this regulator can control the expression of other prophage/OIs in the chromosome of EHEC; in this same microarray. On a cautionary note, such analyses would not be able to differentiate those regulons target by pleiotropic Ler from those potentially

targeted by RgdR; so expression profiling in a $\Delta rgdR$, Δler and $\Delta rgdR\text{-}ler$ double mutant would go some way in helping to delineate the activities of these regulators also.

Interestingly, OI-51 and OI-51 encoded RgdR proteins do not appear to have evolved in typical EPEC strains (although are found in other group B *E. coli* strains), or have perhaps been lost from this pathotype; although RgdR can up-regulate T3S levels in a wild type EPEC strain (E2348/69). Regulation of the LEE is known to differ between EHEC and EPEC where only one promoter is reported to be present in the latter (Elliott *et al.*, 1998; Kendall *et al.*, 2010; Tree *et al.*, 2009). As RgdR was able to stimulate T3S in an EPEC background, this may indicate that RgdR stimulates T3S by binding to a common LEE regulatory DNA region that is present in both strains; and does not require an EHEC specific factor to do this. Intriguingly, atypical EPEC strains are known to harbour variants of RgdR, although these strains do not contain the EAF virulence plasmid found in typical EPEC strains. The significance of this observation if any is unclear, although perhaps worth noting is the P4 ancestry of OI-51. P4 satellite phage are unusual in that they can follow multiple developmental pathways (lysogenic, lytic, or plasmid) and depend on helper P2 phage for these

activities; in particular for virion formation (Briani *et al.*, 2001; Christie and Calendar, 1990; Halling and Calendar, 1990; Liu *et al.*, 1997). Although unsubstantiated, perhaps OI-51 is in some way incompatible with the EAF plasmid of EPEC, or certain EPEC strains do not contain the necessary complement of prophages in their chromosome that would be required for OI-51 to be functional or for it to provide a selective advantage. RgdR induction of T3S in an EPEC background would argue against this, however, this activity was in the absence of the rest of the genes on OI-51, i.e. the complete prophage unit. Perhaps OI-51 and the EAF (pMAR2) or smaller plasmid (pE2348-2) found in EPEC (Iguchi *et al.*, 2009) require the same replication factors; or are incompatible in some other way. Alternatively, it may be the case that as more sequenced EPEC genomes become available, then RgdR proteins will be identified in this specific pathotype.

RgdR proteins have acquired the capacity to control gene expression by integrating into pre-existing regulatory networks present in the host bacterium. However, some very important questions about RgdR still remain un-answered, for example; what induces *rgdR* expression and when; at what concentrations is RgdR normally found in the cell; where does RgdR localise to in the cell, if

indeed it does; and why does RgdR stimulate T3S. QseC is an inter-kingdom QS regulator that governs multiple colonisation phenotypes in EHEC in response to bacterial molecules known as autoinducers (Barba *et al.*, 2005; Clark and Sperandio, 2005; Iyoda *et al.*, 2006; Sperandio *et al.*, 2002; Jobichen *et al.*, 2007). To investigate the possibility that RgdR may be responding to similar signalling molecules, RgdR activity was assessed in a *qseC* deletion cultured under T3S permissive conditions; however, RgdR was able to stimulate T3S levels in this strain. In addition, RgdR was demonstrated to induce T3S in an EHEC strain cultured in non-T3S permissive LB broth, indicating that this regulator does not strictly require T3S specific signals to exert its control; and that RgdR has the capacity to over-ride 'non-permissive' signals and act alongside other factors that can also be expressed in the cell under these conditions to control T3S. Moreover, RgdR also had the capacity to induce T3S levels in low secreting strains of EHEC and increase T3S levels further in high secreting strains; possibly by usurping Ler's own auto-repression at the *LEE1* promoter (Mellies *et al.*, 2008; Yerushalmi *et al.*, 2008). Use of RT-qPCR to determine levels of *rgdR* transcription under differing environmental conditions should give some idea as to how tightly regulated *rgdR* is in the cell. Unfortunately, attempts to localise RgdR were unsuccessful although a variant

that had been engineered to contain a frame shift mutation was unable to stimulate T3S; indicating that RgdR is a functional protein. Further, analysis of cloned RgdR variants from different *E. coli* strains demonstrated that these variants had pronounced effects on other phenotypes independently of their effect on T3S; supporting that they are stably expressed.

As RgdR is a very small protein it was thought that the best chance to look at its localisation would be through the construction of RgdR proteins with tags recognised by high affinity commercial antibodies; so both His-tag and FLAG-tag constructs were made. However, screening of whole cell and supernatant preparations failed to give any detectable signal for the protein, despite the fact that other fusion proteins were detectable with the same anti-6xHis antibody, and by the finding that the His-tag RgdR construct was just as efficient at inducing T3S as the wild type variant. Expression of this His-tagged RgdR clone in an *E. coli* K12 background also failed to give any detectable signal; although induction of RgdR in K12 does repress motility in this background; indicating that RgdR likely represses motility in both pathogenic and non-pathogenic *E. coli* strains via a shared mechanism and does not require EHEC

specific factors for this regulation. Anti-RgdR antibodies were not raised *in vivo* for this research but may provide a valuable tool in future localisation studies. Until RgdR can be localised, one cannot rule out the possibility that RgdR is somehow making it outside of the cell; either via T3S or via some other secretion system (e.g. type IV secretion). RgdRs ability to alter the charge at the surface of the cell could facilitate such an activity and could provide an explanation as to why RgdR stimulates T3S. A more conservative explanation perhaps is that RgdR induces T3S to coordinate the expression and translocation of other effector proteins that are required for pedestal formation and colonisation in the host (Hemrajani *et al.*, 2010; Kenny *et al.*, 1997; Newton *et al.*, 2010).

Horizontally acquired regulators such as the Pch family of regulators are known to coordinate T3S with effector protein secretion, and results from recent studies have suggested that the *pch* genes are the target genes for integrating various environmental signals that control the expression of the LEE genes (Abe *et al.*, 2008; Iyoda and Watanabe, 2004; Porter *et al.*, 2005; Tobe *et al.*, 2005; Yang *et al.*, 2009). Strain to strain variation in LEE gene expression has also been attributed to prophage dense regions of the genome that lie immediately adjacent to the late regions of lambdoid prophage carrying highly conserved *pchA*, *pchB*, *pchC*

and *pchX* alleles (Yang *et al.*, 2009). Moreover, genomic regions adjacent to these *pch* genes are reported to be extraordinarily variable (*pch* adjacent genomic regions/*pch* AGR) and these regions are reported to have substantial effects on patterns of *LEE1* transcription (Yang *et al.*, 2009). This variation in the *pch* AGR has been proposed to serve as a mechanism through which prophage-mediated events adjust virulence gene expression at unlinked loci (Yang *et al.*, 2009). It is perhaps noteworthy then that RgdR is located adjacent to a PerC like regulator (*pchE*) on prophage OI-51; although this could just be a coincidence. RgdR may therefore activate T3S in order to facilitate and coordinate effector protein secretion into host cells; although this is pure speculation at present. No major differences in A/E lesion formation were observed when an OI-51 and *rgdR* deletion strain were compared *in vitro* for this phenotype, however, this is not to say that other effector functions are not impacted on. OI-51 was demonstrated to be important for EHEC O157:H7 colonisation and persistence in ruminants. This could be a direct result of RgdR acting on T3S regulation, or may reflect co-ordinated regulation with other genes on the island. While there are no identified effector proteins on OI-51, bioinformatics analyses show that *z1841* and *ecs1580* harbour a motif consistent with eukaryotic proteins; and RgdR itself contains a putative protein interaction domain (RGD) and N-myristylation

motif; and truncated variants of RgdR harbour a conserved PB1 domain that can be found in proteins associated with intracellular signalling functions in eukaryotic yeast cells (Moscat *et al.*, 2006; Sumimoto *et al.*, 2007; Terasawa *et al.*, 2001). The context of *rgdR* expression now needs to be studied along with the function of other genes in its operon and locally within this genetic region.

Recent studies have shown that LI and LI/II strains of EHEC produce higher levels of Shiga toxins than bovine-associated LII strains (Zhang *et al.*, 2010) and hyper-virulent SNP clade 8 strains have been shown to have higher levels of LEE expression than EHEC O157:H7 strain Sakai from SNP clade 2 (Abu-Ali *et al.*, 2010). It was a possibility that lineage specific variants of RgdR may control T3S to differing degrees and so were examined in this study. Lineage I/II and lineage II variants were slightly more active for T3S stimulation and were more repressive on motility; however the effects of these variants on Shiga toxin expression were not tested. This is something that would be worth testing, either by antibody detection or by qRT-PCR.

6.1.1.4 Future perspective

This research has made a significant contribution to the field by identifying a genetic region in EHEC that is important for ruminant colonisation and by identifying and beginning to characterise a new family of small bacterial regulators that control gene expression in *E. coli*. Importantly, this work has also generated multiple research avenues that can now be pursued. A grant proposal to address the function(s) of OI-51 encoded RgdR proteins in other *E. coli* that present significant economic and public health burdens would surely be well received; including in commensal *E. coli* that may act as a reservoir and mixing pot for the evolution of virulence associated factors that may ultimately be transferred to other bacteria on prophages and other horizontally acquired genetic elements.

References

- Abe, H., Miyahara, A., Oshima, T., Tashiro, K., Ogura, Y., Kuhara, S., Ogasawara, N., Hayashi, T., and Tobe, T.** (2008) Global regulation by horizontally transferred regulators establishes the pathogenicity of *Escherichia coli*. *DNA Res* **15**: 25-38.
- Abraham, J.M., Freitag, C.S., Clements, J.R., and Eisenstein, B.I.** (1985) An invertible element of DNA Controls Phase Variation of Type 1 Fimbriae of *Escherichia coli*. *PNAS* **82**: 5724-5727.
- Abu-Ali, G.S., Ouellette, L.M., Henderson, S.T., Whittam, T.S., and Manning, S.D.** (2010) Differences in adherence and virulence gene expression between two outbreak strains of enterohaemorrhagic *Escherichia coli* O157:H7. *Microbiology* **156**: 408-419.
- Armstrong, G.L., Hollingsworth, J., and Morris, J.G., Jr.** (1996) Emerging foodborne pathogens: *Escherichia coli* O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. *Epidemiol Rev* **18**: 29-51.
- Asadulghani, M., Ogura, Y., Ooka, T., Itoh, T., Sawaguchi, A., Iguchi, A., Nakayama, K., and Hayashi, T.** (2009) The defective prophage pool of *Escherichia coli* O157: prophage-prophage interactions potentiate horizontal transfer of virulence determinants. *PLoS Pathog* **5**: e1000408.
- Ashida, H., Ogawa, M., Mimuro, H., Kobayashi, T., Sanada, T., and Sasakawa, C.** (2011) *Shigella* are versatile mucosal pathogens that circumvent the host innate immune system. *Curr Opin Immunol* **23**: 448-455.
- Barba, J., Bustamante, V.H., Flores-Valdez, M.A., Deng, W., Finlay, B.B., and Puente, J.L.** (2005) A positive regulatory loop controls expression of the locus of enterocyte effacement-encoded regulators Ler and GrlA. *J Bacteriol* **187**: 7918-7930.
- Barnhart, M.M., and Chapman, M.R.** (2006) Curli biogenesis and function. *Annu Rev Microbiol* **60**: 131-147.

Baruch, K., Gur-Arie, L., Nadler, C., Koby, S., Yerushalmi, G., Ben-Neriah, Y., Yogev, O., Shaulian, E., Guttman, C., Zarivach, R., and Rosenshine, I. (2011) Metalloprotease type III effectors that specifically cleave JNK and NF-kappaB. *Embo J* **30**: 221-231.

Baumler, A.J., and Heffron, F. (1995) Identification and sequence analysis of *lpfABCDE*, a putative fimbrial operon of *Salmonella typhimurium*. *J Bacteriol* **177**: 2087-2097.

Ben Nasr, A., Olsen, A., Sjobring, U., Muller-Esterl, W., and Bjorck, L. (1996) Assembly of human contact phase proteins and release of bradykinin at the surface of curli-expressing *Escherichia coli*. *Mol Microbiol* **20**: 927-935.

Benz, I., and Schmidt, M.A. (1989) Cloning and expression of an adhesin (AIDA-I) involved in diffuse adherence of enteropathogenic *Escherichia coli*. *Infect Immun* **57**: 1506-1511.

Berdichevsky, T., Friedberg, D., Nadler, C., Rokney, A., Oppenheim, A., and Rosenshine, I. (2005) Ler is a negative autoregulator of the *LEE1* operon in enteropathogenic *Escherichia coli*. *J Bacteriol* **187**: 349-357.

Besser, R.E., Griffin, P.M., and Slutsker, L. (1999) *Escherichia coli* O157:H7 gastroenteritis and the hemolytic uremic syndrome: an emerging infectious disease. *Annu Rev Med* **50**: 355-367.

Beutin, L. (2006) Emerging enterohaemorrhagic *Escherichia coli*, causes and effects of the rise of a human pathogen. *J Vet Med B Infect Dis Vet Public Health* **53**: 299-305.

Bhatt, S., Romeo, T., and Kalman, D. (2011) Honing the message: post-transcriptional and post-translational control in attaching and effacing pathogens. *Trends Microbiol.* **19**: 217-214.

Blattner, F.R., Plunkett, G., III, Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., and Shao, Y. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-1462.

Blocker, A., Komoriya, K., and Aizawa, S. (2003) Type III secretion systems and bacterial flagella: insights into their function from structural similarities. *Proc Natl Acad Sci U S A* **100**: 3027-3030.

Blomfield, I.C., Vaughn, V., Rest, R.F., and Eisenstein, B.I. (1991) Allelic exchange in *Escherichia coli* using the *Bacillus subtilis* *sacB* gene and a temperature-sensitive pSC101 replicon. *Mol Microbiol* **5**: 1447-1457.

Boerlin, P., Chen, S., Colbourne, J.K., Johnson, R., De Grandis, S., and Gyles, C. (1998) Evolution of enterohemorrhagic *Escherichia coli* hemolysin plasmids and the locus for enterocyte effacement in shiga toxin-producing *E. coli*. *Infect Immun* **66**: 2553-2561.

Boyce, T.G., Swerdlow, D.L., and Griffin, P.M. (1995) *Escherichia coli* O157:H7 and the hemolytic-uremic syndrome. *N Engl J Med* **333**: 364-368.

Briani, F., Deho, G., Forti, F., and Ghisotti, D. (2001) The plasmid status of satellite bacteriophage P4. *Plasmid* **45**: 1-17.

Brunder, W., Karch, H., and Schmidt, H. (2006) Complete sequence of the large virulence plasmid pSFO157 of the sorbitol-fermenting enterohemorrhagic *Escherichia coli* O157:H- strain 3072/96. *Int J Med Microbiol* **296**: 467-474.

Bugarel, M., Beutin, L., Martin, A., Gill, A., and Fach, P. (2010) Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with hemorrhagic colitis and hemolytic uremic syndrome in humans. *Int J Food Microbiol* **142**: 318-329.

Bugarel, M., Martin, A., Fach, P., and Beutin, L. (2011) Virulence gene profiling of enterohemorrhagic (EHEC) and enteropathogenic (EPEC) *Escherichia coli* strains: a basis for molecular risk assessment of typical and atypical EPEC strains. *BMC Microbiol* **11**: 142.

Burland, V., Shao, Y., Perna, N.T., Plunkett, G., Sofia, H.J., and Blattner, F.R. (1998) The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucleic Acids Res* **26**: 4196-4204.

- Bustamante, V.H., Santana, F.J., Calva, E., and Puente, J.L.** (2001) Transcriptional regulation of type III secretion genes in enteropathogenic *Escherichia coli*: Ler antagonizes H-NS-dependent repression. *Mol Microbiol* **39**: 664-678.
- Bustamante, V.H., Villalba, M.I., Garcia-Angulo, V.A., Vazquez, A., Martinez, L.C., Jimenez, R., and Puente, J.L.** (2001) PerC and GrlA independently regulate Ler expression in enteropathogenic *Escherichia coli*. *Mol Microbiol* **82**: 398-415
- Campellone, K.G., and Leong, J.M.** (2003) Tails of two Tirs: actin pedestal formation by enteropathogenic *E. coli* and enterohemorrhagic *E. coli* O157:H7. *Curr Opin Microbiol* **6**: 82-90.
- Campellone, K.G., Robbins, D., and Leong, J.M.** (2004) EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. *Dev Cell* **7**: 217-228.
- Caprioli, A., Morabito, S., Brugere, H., and Oswald, E.** (2005) Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. *Vet Res* **36**: 289-311.
- Castanie-Cornet, M.P., Penfound, T.A., Smith, D., Elliott, J.F., and Foster, J.W.** (1999) Control of acid resistance in *Escherichia coli*. *J Bacteriol* **181**: 3525-3535.
- Chase-Topping, M., Gally, D., Low, C., Matthews, L., and Woolhouse, M.** (2008) Super-shedding and the link between human infection and livestock carriage of *Escherichia coli* O157. *Nat Rev Microbiol* **6**: 904-912.
- Chatterji, D., and Ojha, A.K.** (2001) Revisiting the stringent response, ppGpp and starvation signaling. *Curr Opin Microbiol* **4**: 160-165.
- Chilcott, G.S., and Hughes, K.T.** (2000) Coupling of flagellar gene expression to flagellar assembly in *Salmonella enterica* serovar typhimurium and *Escherichia coli*. *Microbiol Mol Biol Rev* **64**: 694-708.
- Clarke, M.B., and Sperandio, V.** (2005) Transcriptional regulation of flhDC by QseBC and sigma (FliA) in enterohaemorrhagic *Escherichia coli*. *Mol Microbiol* **57**: 1734-1749.

Clarke, S.C. (2001) Diarrhoeagenic *Escherichia coli*-an emerging problem? *Diagn Microbiol Infect Dis* **41**: 93-98.

Clermont, O., Lescat, M., O'Brien, C.L., Gordon, D.M., Tenaillon, O., and Denamur, E. (2008) Evidence for a human-specific *Escherichia coli* clone. *Environ Microbiol* **10**: 1000-1006.

Coia, J.E. (1998) Nosocomial and laboratory-acquired infection with *Escherichia coli* O157. *J Hosp Infect* **40**: 107-113.

Christie, G.E., and Calendar, R. (1990) Interactions between satellite bacteriophage P4 and its helpers. *Annu Rev Genet* **24**: 465-490.

Dame, R.T., Luijsterburg, M.S., Krin, E., Bertin, P.N., Wagner, R., and Wuite, G.J. (2005) DNA bridging: a property shared among H-NS-like proteins. *J Bacteriol* **187**: 1845-1848.

Dean, P., and Kenny, B. (2009) The effector repertoire of enteropathogenic *E. coli*: ganging up on the host cell. *Curr Opin Microbiol* **12**: 101-109.

Deng, W., Li, Y., Vallance, B.A., and Finlay, B.B. (2001) Locus of enterocyte effacement from *Citrobacter rodentium*: sequence analysis and evidence for horizontal transfer among attaching and effacing pathogens. *Infect Immun* **69**: 6323-6335.

Deng, W., Puente, J.L., Gruenheid, S., Li, Y., Vallance, B.A., Vazquez, A., Barba, J., Ibarra, J.A., O'Donnell, P., Metalnikov, P., Ashman, K., Lee, S., Goode, D., Pawson, T., and Finlay, B.B. (2004) Dissecting virulence: systematic and functional analyses of a pathogenicity island. *Proc Natl Acad Sci U S A* **101**: 3597-3602.

DeVinney, R., Gauthier, A., Abe, A., and Finlay, B.B. (1999) Enteropathogenic *Escherichia coli*: a pathogen that inserts its own receptor into host cells. *Cell Mol Life Sci* **55**: 961-976.

Dong, N., Liu, L., and Shao, F. (2010) A bacterial effector targets host DH-PH domain RhoGEFs and antagonizes macrophage phagocytosis. *Embo J* **29**: 1363-1376.

Dorman, C.J. (2004) H-NS: a universal regulator for a dynamic genome. *Nat Rev Microbiol* **2**: 391-400.

Dorman, C.J., and Kane, K.A. (2009) DNA bridging and anti-bridging: a role for bacterial nucleoid-associated proteins in regulating the expression of laterally acquired genes. *FEMS Microbiol Rev* **33**: 587-592.

Dowd, S.E., and Ishizaki, H. (2006) Microarray based comparison of two *Escherichia coli* O157:H7 lineages. *BMC Microbiol* **6**: 30.

D'Souza, S.E., Ginsberg, M.H., and Plow, E.F. (1991) Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *Trends Biochem Sci* **16**: 246-250.

Duriez, P., Clermont, O., Bonacorsi, S., Bingen, E., Chaventre, A., Elion, J., Picard, B., and Denamur, E. (2001) Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* **147**: 1671-1676.

Elliott, S.J., Wainwright, L.A., McDaniel, T.K., Jarvis, K.G., Deng, Y.K., Lai, L.C., McNamara, B.P., Donnenberg, M.S., and Kaper, J.B. (1998) The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol Microbiol* **28**: 1-4.

Elliott, S.J., Sperandio, V., Giron, J.A., Shin, S., Mellies, J.L., Wainwright, L., Hutcheson, S.W., McDaniel, T.K., and Kaper, J.B. (2000) The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect Immun* **68**: 6115-6126.

Erhardt, M., Namba, K., and Hughes, K.T. (2010) Bacterial nanomachines: the flagellum and type III injectisome. *Cold Spring Harb Perspect Biol* **2**: a000299.

Escobar-Paramo, P., Grenet, K., Le Menac'h, A., Rode, L., Salgado, E., Amorin, C., Gouriou, S., Picard, B., Rahimy, M.C., Andremont, A., Denamur, E., and Ruimy, R. (2004) Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl Environ Microbiol* **70**: 5698-5700.

Escobar-Paramo, P., Le Menac'h, A., Le Gall, T., Amorin, C., Gouriou, S., Picard, B., Skurnik, D., and Denamur, E. (2006) Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates. *Environ Microbiol* **8**: 1975-1984.

Emmerson, J.R., Gally, D.L., and Roe, A.J. (2006) Generation of gene deletions and gene replacements in *Escherichia coli* O157:H7 using a temperature sensitive allelic exchange system. *Biol Proced Online* **8**: 153-162.

Feng, P., Lampel, K.A., Karch, H., and Whittam, T.S. (1998) Genotypic and phenotypic changes in the emergence of *Escherichia coli* O157:H7. *J Infect Dis* **177**: 1750-1753.

Flockhart, A. F., Tree, J. J., Xu, X., Karpiyevich, M., McAteer, S. P., Rosenblum, R., Shaw, D. J., Low, C. J., Best, A., Gannon, V., Laing, C., Murphy, K. C., Leong, J. M., Schneiders, T., La Ragione, R., Gally, D. L (2012) Identification of a novel prophage regulator in *Escherichia coli* controlling the expression of type III secretion. *Mol Microbiology* **83**(1): 208-223

Fontaine, A., Arondel, J., and Sansonetti, P.J. (1988) Role of Shiga toxin in the pathogenesis of bacillary dysentery, studied by using a Tox- mutant of *Shigella dysenteriae* 1. *Infect Immun* **56**: 3099-3109.

Fuchs, S., Muhldorfer, I., Donohue-Rolfe, A., Kerenyi, M., Emody, L., Alexiev, R., Nenkov, P., and Hacker, J. (1999) Influence of RecA on *in vivo* virulence and Shiga toxin 2 production in *Escherichia coli* pathogens. *Microb Pathog* **27**: 13-23.

Gally, D. L., T. Rucker, T., and Blomfield, I. C. (1994) The Leucine-responsive regulatory protein binds to the *fim* switch to control phase variation of type 1 fimbriae expression in *Escherichia coli* K-12. *J Bacteriol* **176**: 5665-5672.

Gally, D.L., Leathart, J., and Blomfield, I.C. (1996) Interaction of FimB and FimE with the *fim* switch that controls the phase variation of type 1 fimbriae in *Escherichia coli* K-12. *Mol Microbiol* **21**: 725-738.

Garmendia, J., Frankel, G., and Crepin, V.F. (2005) Enteropathogenic and enterohemorrhagic *Escherichia coli* infections: translocation, translocation, translocation. *Infect Immun* **73**: 2573-2585.

Gauthier, A., Thomas, N.A., and Finlay, B.B. (2003) Bacterial injection machines. *J Biol Chem* **278**: 25273-25276.

Girardeau, J.P., Bertin, Y., and Martin, C. (2009) Genomic analysis of the PAI ICL3 locus in pathogenic LEE-negative Shiga toxin-producing *Escherichia coli* and *Citrobacter rodentium*. *Microbiology* **155**: 1016-1027.

Giron, J.A., Torres, A.G., Freer, E., and Kaper, J.B. (2002) The flagella of enteropathogenic *Escherichia coli* mediate adherence to epithelial cells. *Mol Microbiol* **44**: 361-379.

Goldberg, M.B., Boyko, S.A., Butters, J.R., Stoeber, J.A., Payne, S.M., and Calderwood, S.B. (1992) Characterization of a *Vibrio cholerae* virulence factor homologous to the family of TonB-dependent proteins. *Mol Microbiol* **6**: 2407-2418.

Gordon, D.M., Clermont, O., Tolley, H., and Denamur, E. (2008) Assigning *Escherichia coli* strains to phylogenetic groups: multi-locus sequence typing versus the PCR triplex method. *Environ Microbiol* **10**: 2484-2496

Grant, A.J., Farris, M., Alefounder, P., Williams, P.H., Woodward, M.J., and O'Connor, C.D. (2003) Co-ordination of pathogenicity island expression by the BipA GTPase in enteropathogenic *Escherichia coli* (EPEC). *Mol Microbiol* **48**: 507-521.

Griffin, P.M., and Tauxe, R.V. (1991) The epidemiology of infections caused by *Escherichia coli* O157:H7, other enterohemorrhagic *E. coli*, and the associated hemolytic uremic syndrome. *Epidemiol Rev* **13**: 60-98.

Gruenheid, S., Sekirov, I., Thomas, N.A., Deng, W., O'Donnell, P., Goode, D., Li, Y., Frey, E.A., Brown, N.F., Metalnikov, P., Pawson, T., Ashman, K., and Finlay, B.B. (2004) Identification and characterization of NleA, a non-LEE-encoded type III translocated virulence factor of enterohaemorrhagic *Escherichia coli* O157:H7. *Mol Microbiol* **51**: 1233-1249.

Grys, T.E., Siegel, M.B., Lathem, W.W., and Welch, R.A. (2005) The StcE protease contributes to intimate adherence of enterohemorrhagic *Escherichia coli* O157:H7 to host cells. *Infect Immun* **73**: 1295-1303.

Hacker, J., and Kaper, J.B. (2000) Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54**: 641-679.

Hagberg, L., Engberg, I., Freter, R., Lam, J., Olling, S., and Svanborg Eden, C. (1983) Ascending, unobstructed urinary tract infection in mice caused by pyelonephritogenic *Escherichia coli* of human origin. *Infect Immun* **40**: 273-283.

Halling, C., and Calendar, R. (1990) Bacteriophage P2 ogr and P4 delta genes act independently and are essential for P4 multiplication. *J Bacteriol* **172**: 3549-3558.

Hammar, M., Arnqvist, A., Bian, Z., Olsen, A., and Normark, S. (1995) Expression of two *csg* operons is required for production of fibronectin- and congo red-binding curli polymers in *Escherichia coli* K-12. *Mol Microbiol* **18**: 661-670.

Hansen, A.M., and Kaper, J.B. (2009) Hfq affects the expression of the LEE pathogenicity island in enterohaemorrhagic *Escherichia coli*. *Mol Microbiol* **73**: 446-465.

Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M., and Shinagawa, H. (2001) Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* **8**: 11-22.

Hemrajani, C., Berger, C.N., Robinson, K.S., Marches, O., Mousnier, A., and Frankel, G. (2010) NleH effectors interact with Bax inhibitor-1 to block apoptosis during enteropathogenic *Escherichia coli* infection. *Proc Natl Acad Sci USA* **107**: 3129-3134.

Hendrickson, H. (2009) Order and disorder during *Escherichia coli* divergence. *PLoS Genet* **5**: e1000335.

Holden, N., Totsika, M., Dixon, L., Catherwood, K., and Gally, D.L. (2007) Regulation of P-fimbrial phase variation frequencies in *Escherichia coli* CFT073. *Infect Immun* **75**: 3325-3334.

Huang, L.H., and Syu, W.J. (2008) GrlA of enterohemorrhagic *Escherichia coli* O157:H7 activates *LEE1* by binding to the promoter region. *J Microbiol Immunol Infect* **41**: 9-16.

Hueck, C.J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* **62**: 379-433.

Hughes, D.T., Clarke, M.B., Yamamoto, K., Rasko, D.A., and Sperandio, V. (2009) The QseC adrenergic signaling cascade in Enterohemorrhagic *E. coli* (EHEC). *PLoS Pathog* **5**: e1000553.

Ibarra, J.A., Villalba, M.I., and Puente, J.L. (2003) Identification of the DNA binding sites of PerA, the transcriptional activator of the *bfp* and *per* operons in enteropathogenic *Escherichia coli*. *J Bacteriol* **185**: 2835-2847.

Iguchi, A., Thomson, N.R., Ogura, Y., Saunders, D., Ooka, T., Henderson, I.R., Harris, D., Asadulghani, M., Kurokawa, K., Dean, P., Kenny, B., Quail, M.A., Thurston, S., Dougan, G., Hayashi, T., Parkhill, J., and Frankel, G. (2009) Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *J Bacteriol* **191**: 347-354.

Iida, M., Okamura, N., Yamazaki, M., Yatsuyanagi, J., Kurazono, T., Suzuki, R., Hiruta, N., Isobe, J., Seto, K., Kawano, K., Narimatsu, H., Ratchtrachenchai, O.A., Okabe, N., and Ito, K. (2010) Classification of *perA* sequences and their correlation with autoaggregation in typical enteropathogenic *Escherichia coli* isolates collected in Japan and Thailand. *Microbiol Immunol* **54**: 184-195.

Islam, M.S., Bingle, L.E., Pallen, M.J., and Busby, S.J. (2011) Organization of the *LEE1* operon regulatory region of enterohaemorrhagic *Escherichia coli* O157:H7 and activation by GrlA. *Mol Microbiol* **79**: 468-483.

Iyoda, S., and Watanabe, H. (2004) Positive effects of multiple *pch* genes on expression of the locus of enterocyte effacement genes and adherence of

enterohaemorrhagic *Escherichia coli* O157:H7 to HEp-2 cells. *Microbiology* **150**: 2357-2571.

Iyoda, S., and Watanabe, H. (2005) ClpXP protease controls expression of the type III protein secretion system through regulation of RpoS and GrlR levels in enterohemorrhagic *Escherichia coli*. *J Bacteriol* **187**: 4086-4094.

Iyoda, S., Koizumi, N., Satou, H., Lu, Y., Saitoh, T., Ohnishi, M., and Watanabe, H. (2006) The GrlR-GrlA regulatory system coordinately controls the expression of flagellar and LEE-encoded type III protein secretion systems in enterohemorrhagic *Escherichia coli*. *J Bacteriol* **188**: 5682-5692.

Janion, C. (2008) Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli*. *Int J Biol Sci* **4**: 338-344.

Jarvis, K.G., Giron, J.A., Jerse, A.E., McDaniel, T.K., Donnenberg, M.S., and Kaper, J.B. (1995) Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. *Proc Natl Acad Sci U S A* **92**: 7996-8000.

Jaureguy, F., Landraud, L., Passet, V., Diancourt, L., Frapy, E., Guigon, G., Carbonnelle, E., Lortholary, O., Clermont, O., Denamur, E., Picard, B., Nassif, X., and Brisse, S. (2008) Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics* **9**: 560.

Jimenez, R., Cruz-Migoni, S.B., Huerta-Saquero, A., Bustamante, V.H., and Puente, J.L. (2010) Molecular characterization of GrlA, a specific positive regulator of *ler* expression in enteropathogenic *Escherichia coli*. *J Bacteriol* **18**: 4627-4642.

Jin, Q., Yuan, Z., Xu, J., Wang, Y., Shen, Y., Lu, W., Wang, J., Liu, H., Yang, J., Yang, F., Zhang, X., Zhang, J., Yang, G., Wu, H., Qu, D., Dong, J., Sun, L., Xue, Y., Zhao, A., Gao, Y., Zhu, J., Kan, B., Ding, K., Chen, S., Cheng, H., Yao, Z., He, B., Chen, R., Ma, D., Qiang, B., Wen, Y., Hou, Y., and Yu, J. (2002) Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity through comparison with genomes of *Escherichia coli* K12 and O157. *Nucleic Acids Res* **30**: 4432-4441.

- Jobichen, C., Li, M., Yerushalmi, G., Tan, Y.W., Mok, Y.K., Rosenshine, I., Leung, K.Y., and Sivaraman, J.** (2007) Structure of GrlR and the implication of its EDED motif in mediating the regulation of type III secretion system in EHEC. *PLoS Pathog* **3**: e69.
- Jordan, D.M., Cornick, N., Torres, A.G., Dean-Nystrom, E.A., Kaper, J.B., and Moon, H.W.** (2004) Long polar fimbriae contribute to colonization by *Escherichia coli* O157:H7 *in vivo*. *Infect Immun* **72**: 6168-6171.
- Kanowalchuck, J., Speirs, I., Starvic, S.** (1977) Vero response to cytotoxin of *Escherichia coli*. *Infect. Immun* **18**: 775-779.
- Kaper, J.B., Nataro, J.P., and Mobley, H.L.** (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**: 123-140.
- Karch, H., Tarr, P.I., and Bielaszewska, M.** (2005) Enterohaemorrhagic *Escherichia coli* in human medicine. *Int J Med Microbiol* **295**: 405-418.
- Karmali MA, Petric M, Lim C, Fleming PC, Arbus GS, Lior H.** (1985) The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *J Infect Dis* **151**:775-782.
- Karmali, M.A.** (1989) Infection by verocytotoxin-producing *Escherichia coli*. *Clin Microbiol Rev* **2**: 15-38.
- Karmali, M.A., Mascarenhas, M., Shen, S., Ziebell, K., Johnson, S., Reid-Smith, R., Isaac-Renton, J., Clark, C., Rahn, K., and Kaper, J.B.** (2003) Association of genomic O island 122 of *Escherichia coli* EDL933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. *J Clin Microbiol* **41**: 4930-4940.
- Karmali, M.A.** (2004) Infection by Shiga toxin-producing *Escherichia coli*: an overview. *Mol Biotechnol* **26**: 117-122.
- Kauffman, F.** (1947) The serology of the *coli* group. *J Immunol* **57**: 71-100.
- Kendall, M.M., Rasko, D.A., and Sperandio, V.** (2011) The LysR-type regulator QseA regulates both characterized and putative virulence genes in enterohaemorrhagic *Escherichia coli* O157:H7. *Mol Microbiol* **76**: 1306-1321.

Kenny, B., Abe, A., Stein, M., and Finlay, B.B. (1997) Enteropathogenic *Escherichia coli* protein secretion is induced in response to conditions similar to those in the gastrointestinal tract. *Infect Immun* **65**: 2606-2612.

Kim, J., Nietfeldt, J., and Benson, A.K. (1999) Octamer-based genome scanning distinguishes a unique subpopulation of *Escherichia coli* O157:H7 strains in cattle. *Proc Natl Acad Sci USA* **96**: 13288-13293.

Kim, J., Thanabalasuriar, A., Chaworth-Musters, T., Fromme, J.C., Frey, E.A., Lario, P.I., Metalnikov, P., Rizg, K., Thomas, N.A., Lee, S.F., Hartland, E.L., Hardwidge, P.R., Pawson, T., Strynadka, N.C., Finlay, B.B., Schekman, R., and Gruenheid, S. (2007) The bacterial virulence factor NleA inhibits cellular protein secretion by disrupting mammalian COPII function. *Cell Host Microbe* **2**: 160-171.

Kitagawa, R., Takaya, A., and Yamamoto, T. (2011) Dual Regulatory Pathways of Flagellar Gene Expression by ClpXP Protease in Enterohaemorrhagic *Escherichia coli*. *Microbiology*.

Kostakioti, M., Hadjifrangiskou, M., Pinkner, J.S., and Hultgren, S.J. (2009) QseC-mediated dephosphorylation of QseB is required for expression of genes associated with virulence in uropathogenic *Escherichia coli*. *Mol Microbiol* **73**: 1020-1031.

Kresse, A.U., Rohde, M., and Guzman, C.A. (1999) The EspD protein of enterohemorrhagic *Escherichia coli* is required for the formation of bacterial surface appendages and is incorporated in the cytoplasmic membranes of target cells. *Infect Immun* **67**: 4834-4842.

Laaberki, M.H., Janabi, N., Oswald, E., and Repoila, F. (2006) Concert of regulators to switch on LEE expression in enterohemorrhagic *Escherichia coli* O157:H7: interplay between Ler, GrlA, HNS and RpoS. *Int J Med Microbiol* **296**: 197-210.

Laing, C.R., Buchanan, C., Taboada, E.N., Zhang, Y., Karmali, M.A., Thomas, J.E., and Gannon, V.P. (2009) *In silico* genomic analyses reveal three distinct lineages of *Escherichia coli* O157:H7, one of which is associated with hyper-virulence. *BMC Genomics* **10**: 287.

- La Ragione, R.M., Best, A., Woodward, M.J., and Wales, A.D.** (2009) *Escherichia coli* O157:H7 colonization in small domestic ruminants. *FEMS Microbiol Rev* **33**: 394-410.
- Lease, R.A., and Belfort, M.** (2000) A *trans*-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc Natl Acad Sci U S A* **97**: 9919-9924.
- Leathart, J.B., and Gally, D.L.** (1998) Regulation of type 1 fimbrial expression in uropathogenic *Escherichia coli*: heterogeneity of expression through sequence changes in the fim switch region. *Mol Microbiol* **28**: 371-381.
- Ledesma, M.A., Ochoa, S.A., Cruz, A., Rocha-Ramirez, L.M., Mas-Oliva, J., Eslava, C.A., Giron, J.A., and Xicohtencatl-Cortes, J.** (2010) The hemorrhagic *coli* pilus (HCP) of *Escherichia coli* O157:H7 is an inducer of proinflammatory cytokine secretion in intestinal epithelial cells. *PLoS One* **5**: e12127.
- Lee, J.H., Kim, Y.G., Cho, M.H., Wood, T.K., and Lee, J.** (2011) Transcriptomic analysis for genetic mechanisms of the factors related to biofilm formation in *Escherichia coli* O157:H7. *Curr Microbiol* **62**: 1321-1330.
- Le Gall, T., Clermont, O., Gouriou, S., Picard, B., Nassif, X., Denamur, E., and Tenaillon, O.** (2007) Extraintestinal virulence is a coincidental by-product of commensalism in B2 phylogenetic group *Escherichia coli* strains. *Mol Biol Evol* **24**: 2373-2384.
- Levin, B.R** (1996) The evolution and maintenance of virulence in microparasites. *Emerg Infect Dis* **2**: 93-102.
- Leyton, D.L., Sloan, J., Hill, R.E., Doughty, S., and Hartland, E.L.** (2003) Transfer region of pO113 from enterohemorrhagic *Escherichia coli*: similarity with R64 and identification of a novel plasmid-encoded autotransporter, EpeA. *Infect Immun* **71**: 6307-6319.
- Levine, M.M.** (1987) *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *J Infect Dis* **155**: 377-389.

Li, B., Koch, W.H., and Cebula, T.A. (1997) Detection and characterization of the fimA gene of *Escherichia coli* O157:H7. *Mol Cell Probes* **11**: 397-406.

Liu, T., Renberg, S.K., and Haggard-Ljungquist, E. (1997) Derepression of prophage P2 by satellite phage P4: cloning of the P4 epsilon gene and identification of its product. *J Virol* **71**: 4502-4508.

Lodato, P.B., and Kaper, J.B. (2009) Post-transcriptional processing of the *LEE4* operon in enterohaemorrhagic *Escherichia coli*. *Mol Microbiol* **71**: 273-290.

Low, A.S., Dziva, F., Torres, A.G., Martinez, J.L., Rosser, T., Naylor, S., Spears, K., Holden, N., Mahajan, A., Findlay, J., Sales, J., Smith, D.G., Low, J.C., Stevens, M.P., and Gally, D.L. (2006) Cloning, expression, and characterization of fimbrial operon F9 from enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **74**: 2233-2244.

Luck, S.N., Bennett-Wood, V., Poon, R., Robins-Browne, R.M., and Hartland, E.L. (2005) Invasion of epithelial cells by locus of enterocyte effacement-negative enterohemorrhagic *Escherichia coli*. *Infect Immun* **73**: 3063-3071.

Luck, S.N., Badea, L., Bennett-Wood, V., Robins-Browne, R., and Hartland, E.L. (2006) Contribution of FliC to epithelial cell invasion by enterohemorrhagic *Escherichia coli* O113:H21. *Infect Immun* **74**: 6999-7004.

Macnab, R.M. (2003) How bacteria assemble flagella. *Annu Rev Microbiol* **57**: 77-100.

Mahajan, A., Currie, C.G., Mackie, S., Tree, J., McAteer, S., McKendrick, I., McNeilly, T.N., Roe, A., La Ragione, R.M., Woodward, M.J., Gally, D.L., and Smith, D.G. (2009) An investigation of the expression and adhesin function of H7 flagella in the interaction of *Escherichia coli* O157 : H7 with bovine intestinal epithelium. *Cell Microbiol* **11**: 121-137.

Marches, O., Covarelli, V., Dahan, S., Cougoule, C., Bhatta, P., Frankel, G., and Caron, E. (2008) EspJ of enteropathogenic and enterohaemorrhagic *Escherichia coli* inhibits opsono-phagocytosis. *Cell Microbiol* **10**: 1104-1115.

Martinez-Argudo, I., Sands, C., and Jepson, M.A. (2007) Translocation of enteropathogenic *Escherichia coli* across an *in vitro* M cell model is regulated by its type III secretion system. *Cell Microbiol* **9**: 1538-1546.

Matthews, L., Low, J.C., Gally, D.L., Pearce, M.C., Mellor, D.J., Heesterbeek, J.A., Chase-Topping, M., Naylor, S.W., Shaw, D.J., Reid, S.W., Gunn, G.J., and Woolhouse, M.E. (2006) Heterogeneous shedding of *Escherichia coli* O157 in cattle and its implications for control. *Proc Natl Acad Sci U S A* **103**: 547-552.

Matsushiro, A., Sato, K., Miyamoto, H., Yamamura, T., and Honda, T. (1999) Induction of prophages of enterohemorrhagic *Escherichia coli* O157:H7 with norfloxacin. *J Bacteriol* **181**: 2257-2260.

McDaniel, T.K., Jarvis, K.G., Donnenberg, M.S., and Kaper, J.B. (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. *Proc Natl Acad Sci U S A* **92**: 1664-1668.

McDaniel, T.K., and Kaper, J.B. (1997) A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. *Mol Microbiol* **23**: 399-407.

Mead, P.S., and Griffin, P.M. (1998) *Escherichia coli* O157:H7. *Lancet* **352**: 1207-1212.

Mellies, J.L., Barron, A.M., and Carmona, A.M. (2007) Enteropathogenic and enterohemorrhagic *Escherichia coli* virulence gene regulation. *Infect Immun* **75**: 4199-4210.

Mellies, J.L., Larabee, F.J., Zarr, M.A., Horback, K.L., Lorenzen, E., and Mavor, D. (2008) Ler interdomain linker is essential for anti-silencing activity in enteropathogenic *Escherichia coli*. *Microbiology* **154**: 3624-3638.

Mellies, J.L., Benison, G., McNitt, W., Mavor, D., Boniface, C., and Larabee, F.J. (2011) Ler of Pathogenic *Escherichia coli* forms toroidal protein-DNA complexes. *Microbiology* **157**: 1123-1133.

Michiels, T., Wattiau, P., Brasseur, R., Ruyschaert, J.M., and Cornelis, G. (1990) Secretion of Yop proteins by Yersiniae. *Infect Immun* **58**: 2840-2849.

Mobley, H.L., Green, D.M., Trifillis, A.L., Johnson, D.E., Chippendale, G.R., Lockatell, C.V., Jones, B.D., and Warren, J.W. (1990) Pyelonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect Immun* **58**: 1281-1289.

Moriel, D.G., Bertoldi, I., Spagnuolo, A., Marchi, S., Rosini, R., Nesta, B., Pastorello, I., Corea, V.A., Torricelli, G., Cartocci, E., Savino, S., Scarselli, M., Dobrindt, U., Hacker, J., Tettelin, H., Tallon, L.J., Sullivan, S., Wieler, L.H., Ewers, C., Pickard, D., Dougan, G., Fontana, M.R., Rappuoli, R., Pizza, M., and Serino, L. (2010) Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* **107**: 9072-9077.

Moscat, J., Diaz-Meco, M.T., Albert, A., and Campuzano, S. (2006) Cell signaling and function organized by PB1 domain interactions. *Mol Cell* **23**: 631-640.

Muhldorfer, I., Hacker, J., Keusch, G.T., Acheson, D.W., Tschape, H., Kane, A.V., Ritter, A., Olschlager, T., and Donohue-Rolfe, A. (1996) Regulation of the Shiga-like toxin II operon in *Escherichia coli*. *Infect Immun* **64**: 495-502.

Nadler, C., Baruch, K., Kobi, S., Mills, E., Haviv, G., Farago, M., Alkalay, I., Bartfeld, S., Meyer, T.F., Ben-Neria, Y., and Rosenshine, I. (2010) The type III secretion effector NleE inhibits NF-kappaB activation. *PLoS Pathog* **6**: e1000743.

Nakanishi, N., Abe, H., Ogura, Y., Hayashi, T., Tashiro, K., Kuhara, S., Sugimoto, N., and Tobe, T. (2006) ppGpp with DksA controls gene expression in the locus of enterocyte effacement (LEE) pathogenicity island of enterohaemorrhagic *Escherichia coli* through activation of two virulence regulatory genes. *Mol Microbiol* **61**: 194-205.

Nataro, J.P., and Kaper, J.B. (1998) Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**: 142-201.

Naylor, S.W., Roe, A.J., Nart, P., Spears, K., Smith, D.G., Low, J.C., and Gally, D.L. (2005) *Escherichia coli* O157:H7 forms attaching and effacing lesions at the

terminal rectum of cattle and colonization requires the *LEE4* operon. *Microbiology* **151**: 2773-2781.

Newton, H.J., Sloan, J., Bennett-Wood, V., Adams, L.M., Robins-Browne, R.M., and Hartland, E.L. (2004) Contribution of long polar fimbriae to the virulence of rabbit-specific enteropathogenic *Escherichia coli*. *Infect Immun* **72**: 1230-1239.

Newton, H.J., Pearson, J.S., Badea, L., Kelly, M., Lucas, M., Holloway, G., Wagstaff, K.M., Dunstone, M.A., Sloan, J., Whisstock, J.C., Kaper, J.B., Robins-Browne, R.M., Jans, D.A., Frankel, G., Phillips, A.D., Coulson, B.S., and Hartland, E.L. (2010) The type III effectors NleE and NleB from enteropathogenic *E. coli* and OspZ from *Shigella* block nuclear translocation of NF-kappaB p65. *PLoS Pathog* **6**: e1000898.

Nie, H., Yang, F., Zhang, X., Yang, J., Chen, L., Wang, J., Xiong, Z., Peng, J., Sun, L., Dong, J., Xue, Y., Xu, X., Chen, S., Yao, Z., Shen, Y., and Jin, Q. (2006) Complete genome sequence of *Shigella flexneri* 5b and comparison with *Shigella flexneri* 2a. *BMC Genomics* **7**: 173.

Nieto, J.M., Madrid, C., Prenafeta, A., Miquelay, E., Balsalobre, C., Carrascal, M., and Juarez, A. (2000) Expression of the hemolysin operon in *Escherichia coli* is modulated by a nucleoid-protein complex that includes the proteins Hha and H-NS. *Mol Gen Genet* **263**: 349-358.

Ochman, H., and Selander, R.K. (1984) Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* **157**: 690-693.

Ogura, Y., Ooka, T., Iguchi, A., Toh, H., Asadulghani, M., Oshima, K., Kodama, T., Abe, H., Nakayama, K., Kurokawa, K., Tobe, T., Hattori, M., and Hayashi, T. (2009) Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proc Natl Acad Sci USA* **106**: 17939-17944.

Ohnishi, M., and Hayashi, T. (2002) Genetic diversity of enterohemorrhagic *Escherichia coli*]. *Nippon Rinsho* **60**: 1077-1082.

Olsen, A., Arnqvist, A., Hammar, M., Sukupolvi, S., and Normark, S. (1993) The RpoS sigma factor relieves H-NS-mediated transcriptional repression of

csgA, the subunit gene of fibronectin-binding curli in *Escherichia coli*. *Mol Microbiol* **7**: 523-536.

Osawa, R., Iyoda, S., Nakayama, S.I., Wada, A., Yamai, S., and Watanabe, H. (2000) Genotypic variations of Shiga toxin-converting phages from enterohaemorrhagic *Escherichia coli* O157: H7 isolates. *J Med Microbiol* **49**: 565-574.

Paton, A.W., and Paton, J.C. (1998) Detection and characterization of Shiga toxigenic *Escherichia coli* by using multiplex PCR assays for *stx1*, *stx2*, *eaeA*, enterohemorrhagic *E. coli hlyA*, *rfbO111*, and *rfbO157*. *J Clin Microbiol* **36**: 598-602.

Payne, S.M., and Finkelstein, R.A. (1977) Detection and differentiation of iron-responsive avirulent mutants on Congo red agar. *Infect Immun* **18**: 94-98.

Perna, N.T., Plunkett, G., Burland, V., Mau, B., Glasner, J.D., Rose, D.J., Mayhew, G.F., Evans, P.S., Gregor, J., Kirkpatrick, H.A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E.J., Davis, N.W., Lim, A., Dimalanta, E.T., Potamousis, K.D., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A., and Blattner, F.R. (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**: 529-533.

Plunkett, G., 3rd, Rose, D.J., Durfee, T.J., and Blattner, F.R. (1999) Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J Bacteriol* **181**: 1767-1778.

Pore, D., Mahata, N., Pal, A., and Chakrabarti, M.K. (2011) Outer membrane protein A (OmpA) of *Shigella flexneri* 2a, induces protective immune response in a mouse model. *PLoS One* **6**: e22663.

Porter, M.E., Mitchell, P., Roe, A.J., Free, A., Smith, D.G., and Gally, D.L. (2004) Direct and indirect transcriptional activation of virulence genes by an AraC-like protein, PerA from enteropathogenic *Escherichia coli*. *Mol Microbiol* **54**: 1117-1133.

Porter, M.E., Mitchell, P., Free, A., Smith, D.G., and Gally, D.L. (2005) The *LEE1* promoters from both enteropathogenic and enterohemorrhagic *Escherichia*

coli can be activated by PerC-like proteins from either organism. *J Bacteriol* **187**: 458-472.

Prigent-Combaret, C., Brombacher, E., Vidal, O., Ambert, A., Lejeune, P., Landini, P., and Dorel, C. (2001) Complex regulatory network controls initial adhesion and biofilm formation in *Escherichia coli* via regulation of the *csgD* gene. *J Bacteriol* **183**: 7213-7223.

Rashid, R.A., Tabata, T.A., Oatley, M.J., Besser, T.E., Tarr, P.I., and Moseley, S.L. (2006) Expression of putative virulence factors of *Escherichia coli* O157:H7 differs in bovine and human infections. *Infect Immun* **74**: 4142-4148.

Rasko, D., Rosovitz, M., Myers, G., Seshadri, R., Cer, R., Jiang, L., Ravel, J., Fricke, W.F. and Sebastian, Y (2008) Direct submission (NCBI). J. Craig Venter Institute, 9704 Medical Centre, Rockville, MD 20850, USA.

Reading, N.C., Torres, A.G., Kendall, M.M., Hughes, D.T., Yamamoto, K., and Sperandio, V. (2007) A novel two-component signaling system that activates transcription of an enterohemorrhagic *Escherichia coli* effector involved in remodeling of host actin. *J Bacteriol* **189**: 2468-2476.

Reading, N.C., Rasko, D.A., Torres, A.G., and Sperandio, V. (2009) The two-component system QseEF and the membrane protein QseG link adrenergic and stress sensing to bacterial pathogenesis. *Proc Natl Acad Sci U S A* **106**: 5889-5894.

Riley LW, Remis RS, Helgerson SD, McGee HB, Wells JG, Davis BR, Hebert RJ, Olcott ES, Johnson LM, Hargrett NT, Blake PA, Cohen ML. (1983) Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N Engl J Med* **308**:681-685.

Robins-Browne, R.M., and Hartland, E.L. (2002) *Escherichia coli* as a cause of diarrhea. *J Gastroenterol Hepatol* **17**: 467-475.

Roe, A.J., Currie, C., Smith, D.G., and Gally, D.L. (2001) Analysis of type 1 fimbriae expression in verotoxigenic *Escherichia coli*: a comparison between serotypes O157 and O26. *Microbiology* **147**: 145-152.

Roe, A.J., Hoey, D.E., and Gally, D.L. (2003) Regulation, secretion and activity of type III-secreted proteins of enterohaemorrhagic *Escherichia coli* O157. *Biochem Soc Trans* **31**: 98-103.

Roe, A.J., Naylor, S.W., Spears, K.J., Yull, H.M., Dransfield, T.A., Oxford, M., McKendrick, I.J., Porter, M., Woodward, M.J., Smith, D.G., and Gally, D.L. (2004) Co-ordinate single-cell expression of *LEE4*- and *LEE5*-encoded proteins of *Escherichia coli* O157:H7. *Mol Microbiol* **54**: 337-352.

Romeo, T. (1998) Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Mol Microbiol* **29**: 1321-1330.

Royan SV, Jones RM, Koutsouris A, Roxas JL, Falzari K, Weflen AW, Kim A, Bellmeyer A, Turner JR, Neish AS, Rhee KJ, Viswanathan VK, Hecht GA. (2010) Enteropathogenic *E. coli* non-LEE encoded effectors NleH1 and NleH2 attenuate NF- κ B activation. *Mol. Microbiol* **5**: 1232-1245.

Ruoslahti, E. (1996) RGD and other recognition sequences for integrins. *Annu Rev Cell Dev Biol* **12**: 697-715.

Saitoh, T., Iyoda, S., Yamamoto, S., Lu, Y., Shimuta, K., Ohnishi, M., Terajima, J., and Watanabe, H. (2008) Transcription of the *ehx* enterohemolysin gene is positively regulated by GrlA, a global regulator encoded within the locus of enterocyte effacement in enterohemorrhagic *Escherichia coli*. *J Bacteriol* **190**: 4822-4830.

Schmidt, H., Beutin, L., and Karch, H. (1995) Molecular analysis of the plasmid-encoded hemolysin of *Escherichia coli* O157:H7 strain EDL 933. *Infect Immun* **63**: 1055-1061.

Schmidt, M.A. (2010) LEEways: tales of EPEC, ATEC and EPEC. *Cell Microbiol* **12**: 1544-52.

Schneiders, T., and Levy, S.B. (2006) MarA-mediated transcriptional repression of the *rob* promoter. *J Biol Chem* **281**: 10049-10055.

Schroeder, G.N., and Hilbi, H. (2008) Molecular pathogenesis of *Shigella* spp.: controlling host cell signaling, invasion, and death by type III secretion. *Clin Microbiol Rev* **21**: 134-156.

- Schwidder, M., Hensel, M., and Schmidt, H.** (2011) Regulation of *nleA* in Shiga Toxin-Producing *Escherichia coli* O84:H4 Strain 4795/97. *J Bacteriol* **193**: 832-841.
- Shakhnovich, E.A., Davis, B.M., and Waldor, M.K.** (2009) Hfq negatively regulates type III secretion in EHEC and several other pathogens. *Mol Microbiol* **74**: 347-363.
- Sharma, V.K., and Zuerner, R.L.** (2004) Role of *hha* and *ler* in transcriptional regulation of the *esp* operon of enterohemorrhagic *Escherichia coli* O157:H7. *J Bacteriol* **186**: 7290-7301.
- Sharp, F.C., and Sperandio, V.** (2007) QseA directly activates transcription of *LEE1* in enterohemorrhagic *Escherichia coli*. *Infect Immun* **75**: 2432-2440.
- Sjobering, U., Pohl, G., and Olsen, A.** (1994) Plasminogen, absorbed by *Escherichia coli* expressing curli or by *Salmonella enteritidis* expressing thin aggregative fimbriae, can be activated by simultaneously captured tissue-type plasminogen activator (t-PA). *Mol Microbiol* **14**: 443-452.
- Small, P., Blankenhorn, D., Welty, D., Zinser, E., and Slonczewski, J.L.** (1994) Acid and base resistance in *Escherichia coli* and *Shigella flexneri*: role of *rpoS* and growth pH. *J Bacteriol* **176**: 1729-1737.
- Sperandio, V., Mellies, J.L., Nguyen, W., Shin, S., and Kaper, J.B.** (1999) Quorum sensing controls expression of the type III secretion gene transcription and protein secretion in enterohemorrhagic and enteropathogenic *Escherichia coli*. *Proc Natl Acad Sci U S A* **96**: 15196-15201.
- Sperandio, V., Torres, A.G., Giron, J.A., and Kaper, J.B.** (2001) Quorum sensing is a global regulatory mechanism in enterohemorrhagic *Escherichia coli* O157:H7. *J Bacteriol* **183**: 5187-5197.
- Sperandio, V., Torres, A.G., Jarvis, B., Nataro, J.P., and Kaper, J.B.** (2003) Bacteria-host communication: the language of hormones. *Proc Natl Acad Sci U S A* **100**: 8951-8956.
- Steele, M., Ziebell, K., Zhang, Y., Benson, A., Konczy, P., Johnson, R., and Gannon, V.** (2007) Identification of *Escherichia coli* O157:H7 genomic regions

conserved in strains with a genotype associated with human infection. *Appl Environ Microbiol* **73**: 22-31.

Steele, M., Ziebell, K., Zhang, Y., Benson, A., Konczy, P., Johnson, R., and Gannon, V. (2007) Identification of *Escherichia coli* O157:H7 genomic regions conserved in strains with a genotype associated with human infection. *Appl Environ Microbiol* **73**: 22-31.

Spears, K.J., Roe, A.J., and Gally, D.L. (2006) A comparison of enteropathogenic and enterohaemorrhagic *Escherichia coli* pathogenesis. *FEMS Microbiol Lett* **255**: 187-202.

Sperandio, V., Li, C.C., and Kaper, J.B. (2002) Quorum-sensing *Escherichia coli* regulator A: a regulator of the LysR family involved in the regulation of the locus of enterocyte effacement pathogenicity island in enterohemorrhagic *E. coli*. *Infect Immun* **70**: 3085-3093.

Son, W.G., Graham, T.A., and Gannon, V.P. (2002) Immunological characterization of *Escherichia coli* O157:H7 intimin gamma1. *Clin Diagn Lab Immunol* **9**: 46-53.

Stoebel, D.M., Free, A., and Dorman, C.J. (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology* **154**: 2533-2545.

Strachan, N.J., Dunn, G.M., Locking, M.E., Reid, T.M., and Ogden, I.D. (2006) *Escherichia coli* O157: burger bug or environmental pathogen? *Int J Food Microbiol* **112**: 129-137.

Su, C., and Brandt, L.J. (1995) *Escherichia coli* O157:H7 infection in humans. *Ann Intern Med* **123**: 698-714.

Sumimoto, H., Kamakura, S., and Ito, T. (2007) Structure and function of the PB1 domain, a protein interaction module conserved in animals, fungi, amoebas, and plants. *Sci STKE* **2007**: re6.

Tarr, P.I., Bilge, S.S., Vary, J.C., Jr., Jelacic, S., Habeeb, R.L., Ward, T.R., Baylor, M.R., and Besser, T.E. (2000) Iha: a novel *Escherichia coli* O157:H7

adherence-conferring molecule encoded on a recently acquired chromosomal island of conserved structure. *Infect Immun* **68**: 1400-1407.

Tatsuno, I., Nagano, K., Taguchi, K., Rong, L., Mori, H., and Sasakawa, C. (2003) Increased adherence to Caco-2 cells caused by disruption of the *yhiE* and *yhiF* genes in enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **71**: 2598-2606.

Taylor, D.E., Rooker, M., Keelan, M., Ng, L.K., Martin, I., Perna, N.T., Burland, N.T., and Blattner, F.R. (2002) Genomic variability of O islands encoding tellurite resistance in enterohemorrhagic *Escherichia coli* O157:H7 isolates. *J Bacteriol* **184**: 4690-4698.

Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010) The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* **8**: 207-217.

Terasawa, H., Noda, Y., Ito, T., Hatanaka, H., Ichikawa, S., Ogura, K., Sumimoto, H., and Inagaki, F. (2001) Structure and ligand recognition of the PB1 domain: a novel protein module binding to the PC motif. *EMBO J* **20**: 3947-3956.

Thanabalasuriar, A., Koutsouris, A., Weflen, A., Mimee, M., Hecht, G., and Gruenheid, S. (2010) The bacterial virulence factor NleA is required for the disruption of intestinal tight junctions by enteropathogenic *Escherichia coli*. *Cell Microbiol* **12**: 31-41.

Timmermans, J., and Van Melderren, L. (2010) Post-transcriptional global regulation by CsrA in bacteria. *Cell Mol Life Sci* **67**: 2897-2908.

Tobe, T., Ando, H., Ishikawa, H., Abe, H., Tashiro, K., Hayashi, T., Kuhara, S., and Sugimoto, N. (2005) Dual regulatory pathways integrating the RcsC-RcsD-RcsB signalling system control enterohaemorrhagic *Escherichia coli* pathogenicity. *Mol Microbiol* **58**: 320-333.

Tobe, T., Beatson, S.A., Taniguchi, H., Abe, H., Bailey, C.M., Fivian, A., Younis, R., Matthews, S., Marches, O., Frankel, G., Hayashi, T., and Pallen, M.J. (2006) An extensive repertoire of type III secretion effectors in *Escherichia*

coli O157 and the role of lambdoid phages in their dissemination. *Proc Natl Acad Sci U S A* **103**: 14941-14946.

Tomoyasu, T., Takaya, A., Handa, Y., Karata, K., and Yamamoto, T. (2005) ClpXP controls the expression of LEE genes in enterohaemorrhagic *Escherichia coli*. *FEMS Microbiol Lett* **253**: 59-66.

Torres, A.G., Giron, J.A., Perna, N.T., Burland, V., Blattner, F.R., Avelino-Flores, F., and Kaper, J.B. (2002) Identification and characterization of *lpfABCC'DE*, a fimbrial operon of enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **70**: 5416-5427.

Torres, A.G., Milflores-Flores, L., Garcia-Gallegos, J.G., Patel, S.D., Best, A., La Ragione, R.M., Martinez-Laguna, Y., and Woodward, M.J. (2007) Environmental regulation and colonization attributes of the long polar fimbriae (LPF) of *Escherichia coli* O157:H7. *Int J Med Microbiol* **297**: 177-185.

Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, M.E., Frapy, E., Garry, L., Ghigo, J.M., Gilles, A.M., Johnson, J., Le Bouguenec, C., Lescat, M., Mangenot, S., Martinez-Jehanne, V., Matic, I., Nassif, X., Oztas, S., Petit, M.A., Pichon, C., Rouy, Z., Ruf, C.S., Schneider, D., Turret, J., Vacherie, B., Vallenet, D., Medigue, C., Rocha, E.P., and Denamur, E. (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* **5**: e1000344.

Tozzoli, R., Caprioli, A., and Morabito, S. (2005) Detection of *toxB*, a plasmid virulence gene of *Escherichia coli* O157, in enterohemorrhagic and enteropathogenic *E. coli*. *J Clin Microbiol* **43**: 4052-4056.

Tree, J.J., Wolfson, E.B., Wang, D., Roe, A.J., and Gally, D.L. (2009) Controlling injection: regulation of type III secretion in enterohaemorrhagic *Escherichia coli*. *Trends Microbiol* **17**: 361-370.

Tree, J.J., Roe, A.J., Flockhart, A., McAteer, S.P., Xu, X., Shaw, D., Mahajan, A., Beatson, S.A., Best, A., Lotz, S., Woodward, M.J., La Ragione, R., Murphy, K.C., Leong, J.M., and Gally, D.L. (2011) Transcriptional regulators of the GAD

acid stress island are carried by effector protein-encoding prophages and indirectly control type III secretion in enterohemorrhagic *Escherichia coli* O157:H7. *Mol Microbiol* **80**: 1349-1365.

Tsui, H.C., Leung, H.C., and Winkler, M.E. (1994) Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12. *Mol Microbiol* **13**: 35-49.

Tupper, A.E., Owen-Hughes, T.A., Ussery, D.W., Santos, D.S., Ferguson, D.J., Sidebotham, J.M., Hinton, J.C., and Higgins, C.F. (1994) The chromatin-associated protein H-NS alters DNA topology *in vitro*. *EMBO J* **13**: 258-268.

Uhlich, G.A., Keen, J.E., and Elder, R.O. (2001) Mutations in the *csgD* promoter associated with variations in curli expression in certain strains of *Escherichia coli* O157:H7. *Appl Environ Microbiol* **67**: 2367-2370.

Umanski, T., Rosenshine, I., and Friedberg, D. (2002) Thermoregulated expression of virulence genes in enteropathogenic *Escherichia coli*. *Microbiology* **148**: 2735-2744.

Vossenkamper, A., Macdonald, T.T., and Marches, O. (2011) Always one step ahead: How pathogenic bacteria use the type III secretion system to manipulate the intestinal mucosal immune system. *J Inflamm (Lond)* **8**: 11.

Wales, A.D., Clifton-Hadley, F.A., Cookson, A.L., Dibb-Fuller, M.P., La Ragione, R.M., Sprigings, K.A., Pearson, G.R., and Woodward, M.J. (2001) Experimental infection of six-month-old sheep with *Escherichia coli* O157:H7. *Vet Rec* **148**: 630-631.

Wang, R.F., and Kushner, S.R. (1991) Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **100**: 195-199.

Wang, X., Dubey, A.K., Suzuki, K., Baker, C.S., Babitzke, P., and Romeo, T. (2005) CsrA post-transcriptionally represses *pgaABCD*, responsible for synthesis of a biofilm polysaccharide adhesin of *Escherichia coli*. *Mol Microbiol* **56**: 1648-1663.

- Weflen, A.W., Alto, N.M., Viswanathan, V.K., and Hecht, G.** (2010) *E. coli* secreted protein F promotes EPEC invasion of intestinal epithelial cells via an SNX9-dependent mechanism. *Cell Microbiol* **12**: 919-929.
- Wei, Y., Vollmer, A.C., and LaRossa, R.A.** (2001) *In vivo* titration of mitomycin C action by four *Escherichia coli* genomic regions on multicopy plasmids. *J Bacteriol* **183**: 2259-2264.
- Wei, J., Goldberg, M.B., Burland, V., Venkatesan, M.M., Deng, W., Fournier, G., Mayhew, G.F., Plunkett, G., 3rd, Rose, D.J., Darling, A., Mau, B., Perna, N.T., Payne, S.M., Runyen-Janecky, L.J., Zhou, S., Schwartz, D.C., and Blattner, F.R.** (2003) Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. *Infect Immun* **71**: 2775-2786.
- Welch, R.A., Burland, V., Plunkett, G., 3rd, Redford, P., Roesch, P., Rasko, D., Buckles, E.L., Liou, S.R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C., Perna, N.T., Mobley, H.L., Donnenberg, M.S., and Blattner, F.R.** (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* **99**: 17020-17024.
- Wells, T.J., Sherlock, O., Rivas, L., Mahajan, A., Beatson, S.A., Torpdahl, M., Webb, R.I., Allsopp, L.P., Gobius, K.S., Gally, D.L., and Schembri, M.A.** (2008) EhaA is a novel autotransporter protein of enterohemorrhagic *Escherichia coli* O157:H7 that contributes to adhesion and biofilm formation. *Environ Microbiol* **10**: 589-604.
- Whittam, T.S., Wolfe, M.L., Wachsmuth, I.K., Orskov, F., Orskov, I., and Wilson, R.A.** (1993) Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. *Infect Immun* **61**: 1619-1629.
- Whittam, T.S., and Bumbaugh, A.C.** (2002) Inferences from whole-genome sequences of bacterial pathogens. *Curr Opin Genet Dev* **12**: 719-725.
- Wieler, L.H., McDaniel, T.K., Whittam, T.S., and Kaper, J.B.** (1997) Insertion site of the locus of enterocyte effacement in enteropathogenic and enterohemorrhagic *Escherichia coli* differs in relation to the clonal phylogeny of the strains. *FEMS Microbiol Lett* **156**: 49-53.

Xicohtencatl-Cortes, J., Monteiro-Neto, V., Ledesma, M.A., Jordan, D.M., Francetic, O., Kaper, J.B., Puente, J.L., and Giron, J.A. (2007) Intestinal adherence associated with type IV pili of enterohemorrhagic *Escherichia coli* O157:H7. *J Clin Invest* **117**: 3519-3529.

Yang, Z., Kim, J., Zhang, C., Zhang, M., Nietfeldt, J., Southward, C.M., Surette, M.G., Kachman, S.D., and Benson, A.K. (2009) Genomic instability in regions adjacent to a highly conserved pch prophage in *Escherichia coli* O157:H7 generates diversity in expression patterns of the LEE pathogenicity island. *J Bacteriol* **191**: 3553-3568.

Ye, C., Lan, R., Xia, S., Zhang, J., Sun, Q., Zhang, S., Jing, H., Wang, L., Li, Z., Zhou, Z., Zhao, A., Cui, Z., Cao, J., Jin, D., Huang, L., Wang, Y., Luo, X., Bai, X., Wang, Y., Wang, P., Xu, Q., and Xu, J. (2010) Emergence of a new multidrug-resistant serotype X variant in an epidemic clone of *Shigella flexneri*. *J Clin Microbiol* **48**: 419-426.

Yen, H., Ooka, T., Iguchi, A., Hayashi, T., Sugimoto, N., and Tobe, T. (2010) NleC, a type III secretion protease, compromises NF-kappaB activation by targeting p65/RelA. *PLoS Pathog* **6**: e1001231.

Yerushalmi, G., Nadler, C., Berdichevski, T., and Rosenshine, I. (2008) Mutational analysis of the locus of enterocyte effacement-encoded regulator (Ler) of enteropathogenic *Escherichia coli*. *J Bacteriol* **190**: 7808-7818.

Yin, X., Chambers, J.R., Wheatcroft, R., Johnson, R.P., Zhu, J., Liu, B., and Gyles, C.L. (2009) Adherence of *Escherichia coli* O157:H7 mutants *in vitro* and in ligated pig intestines. *Appl Environ Microbiol* **75**: 4975-4983.

Yona-Nadler, C., Umanski, T., Aizawa, S., Friedberg, D., and Rosenshine, I. (2003) Integration host factor (IHF) mediates repression of flagella in enteropathogenic and enterohaemorrhagic *Escherichia coli*. *Microbiology* **149**: 877-884.

Zhang, L., Chaudhuri, R.R., Constantinidou, C., Hobman, J.L., Patel, M.D., Jones, A.C., Sarti, D., Roe, A.J., Vlisidou, I., Shaw, R.K., Falciani, F., Stevens, M.P., Gally, D.L., Knutton, S., Frankel, G., Penn, C.W., and Pallen, M.J. (2004) Regulators encoded in the *Escherichia coli* type III secretion system 2 gene cluster

influence expression of genes within the locus for enterocyte effacement in enterohemorrhagic *E. coli* O157:H7. *Infect Immun* **72**: 7282-7293.

Zhang, W., Qi, W., Albert, T.J., Motiwala, A.S., Alland, D., Hyytia-Trees, E.K., Ribot, E.M., Fields, P.I., Whittam, T.S., and Swaminathan, B. (2006) Probing genomic diversity and evolution of *Escherichia coli* O157 by single nucleotide polymorphisms. *Genome Res* **16**: 757-767.

Zhang, Y., Laing, C., Steele, M., Ziebell, K., Johnson, R., Benson, A.K., Taboada, E., and Gannon, V.P. (2007) Genome evolution in major *Escherichia coli* O157:H7 lineages. *BMC Genomics* **8**: 121.

Zhang, Y., C. Laing, Z. Zhang, J. Hallewell, C. You, K. Ziebell, R. P. Johnson, A.M. Kropinski, J.E. Thomas, M. Karmali and Gannon V.P.J. (2010) Lineage and host-source are both correlated with levels of Shiga toxin 2 Production by *Escherichia coli* O157:H7 strains. *App Env Microbiol* **76**:474-82.

Zhou, Z., Li, X., Liu, B., Beutin, L., Xu, J., Ren, Y., Feng, L., Lan, R., Reeves, P.R., and Wang, L. (2010) Derivation of *Escherichia coli* O157:H7 from its O55:H7 precursor. *PLoS One* **5**: e8700.

Zhu, C., Agin, T.S., Elliott, S.J., Johnson, L.A., Thate, T.E., Kaper, J.B., and Boedeker, E.C. (2001) Complete nucleotide sequence and analysis of the locus of enterocyte effacement from rabbit diarrheagenic *Escherichia coli* RDEC-1. *Infect Immun* **69**: 2107-2115.

Identification of a novel prophage regulator in *Escherichia coli* controlling the expression of type III secretion

Allen F. Flockhart,¹ Jai J. Tree,¹ Xuefang Xu,¹ Maryia Karpiyevich,¹ Sean P. McAteer,¹ Ronen Rosenblum,² Darren J. Shaw,^{1,3} Christopher J. Low,³ Angus Best,⁴ Victor Gannon,⁵ Chad Laing,⁵ Kenan C. Murphy,⁶ John M. Leong,⁶ Thamarai Schneiders,² Roberto La Ragione^{4,7} and David L. Gally^{1*}

¹Immunity and Infection Division, The Roslin Institute and R(D)SVS, University of Edinburgh, Edinburgh EH25 9RG, UK.

²Centre for infection and Immunity, 4th Floor Medical Biology Centre, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, Belfast BT9 7BL, UK.

³Division of Veterinary Clinical Sciences, Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Easter Bush Campus, Roslin, Midlothian EH25 9RG, UK.

⁴Department of Food and Environmental Safety, Veterinary Laboratories Agency, Woodham Lane, New Haw, Addlestone, Surrey, KT15 3NB, UK.

⁵Laboratory for Foodborne Zoonoses, Public Health Agency of Canada, 1st floor, C.F.I.A. Building, Lethbridge, AB T1J 3Z4, Canada.

⁶Department of Molecular Genetics and Microbiology, University of Massachusetts Medical School, Worcester, MA 01655, USA.

⁷Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK.

Summary

This study has identified horizontally acquired genomic regions of enterohaemorrhagic *Escherichia coli* O157:H7 that regulate expression of the type III secretion (T3S) system encoded by the locus of enterocyte effacement (LEE). Deletion of O-island 51, a 14.93 kb cryptic prophage (CP-933C), resulted in a

reduction in LEE expression and T3S. The deletion also had a reduced capacity to attach to epithelial cells and significantly reduced *E. coli* O157 excretion levels from sheep. Further characterization of O-island 51 identified a novel positive regulator of the LEE, encoded by *ecs1581* in the *E. coli* O157:H7 strain Sakai genome and present but not annotated in the *E. coli* strain EDL933 sequence. Functionally important residues of *ECs1581* were identified based on phenotypic variants present in sequenced *E. coli* strains and the regulator was termed RgdR based on a motif demonstrated to be important for stimulation of gene expression. While RgdR activated expression from the LEE1 promoter in the presence or absence of the LEE-encoded regulator (Ler), RgdR stimulation of T3S required *ler* and Ler autoregulation. RgdR also controlled the expression of other phenotypes, including motility, indicating that this new family of regulators may have a more global role in *E. coli* gene expression.

Introduction

Escherichia coli strains are usually present in the flora of mammalian gastrointestinal (GI) tracts and many are considered non-pathogenic. However, some strains are associated with serious intestinal and extra-intestinal infections. The main differences among strains of these different pathotypes can be attributed to the acquisition of genetic information from mobile genetic elements, in particular bacteriophage (Kaper *et al.*, 2004). Bacteriophage integration and recombination not only leads to phage evolution, but is a key driver in the acquisition of virulence attributes in enteric bacteria, and of the complex regulatory networks that control their expression. Enterohaemorrhagic *E. coli* (EHEC) contain prophage-encoded Shiga toxins and are associated with severe GI and systemic disease in humans (Nataro and Kaper, 1998; Karmali, 2004). Ruminants are considered to be the most important reservoirs for EHEC, particularly cattle and sheep which shed the organism in their faeces (La Ragione *et al.*, 2009). The predominant pathogenic serotype in North America, parts of Asia and the UK is O157:H7 (Armstrong *et al.*, 1996; Besser *et al.*, 1999; Caprioli *et al.*, 2005).

Studies on EHEC O157:H7 have shown that the genotypic diversity among different strains is largely attributable to bacteriophage-related DNA fragments (Ohnishi and Hayashi, 2002; Zhang *et al.*, 2007; Asadulghani *et al.*,

Accepted 14 November, 2011. *For correspondence. E-mail dgally@ed.ac.uk; Tel. (+44) 131 6519242; Fax (+44) 131 6519105.

This article Identification of a novel prophage regulator in *Escherichia coli* controlling the expression of type III secretion was written by Roberto La Ragione of Animal Health and Veterinary Laboratories Agency. It is published with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

Re-use of this article is permitted in accordance with the Terms and Conditions set out at http://wileyonlinelibrary.com/onlineopen/OnlineOpen_Terms

2009). In EHEC O157:H7 strain EDL933, these horizontally acquired elements have been termed O-islands (OIs) and include both fully functional and cryptic prophages (Perna *et al.*, 2001). The locus of enterocyte effacement (LEE) encodes a type III secretion (T3S) system, an essential virulence factor for EHEC that allows the direct injection of bacterial effector proteins into host cells to subvert host cell signalling pathways and promote bacterial attachment (Kenny *et al.*, 1997; Hemrajani *et al.*, 2010; Newton *et al.*, 2010). While T3S has been shown to be critical for EHEC O157:H7 colonization in the ruminant host (Naylor *et al.*, 2005), levels of T3S can vary significantly among different EHEC isolates (Son *et al.*, 2002; Roe *et al.*, 2003; Rashid *et al.*, 2006; Yang *et al.*, 2009). Such variation in T3S regulation may have important implications for bacterial persistence and shedding in the bovine host and for human infection (Chase-Topping *et al.*, 2008). Recent studies comparing EHEC O157:H7 genome sequences have identified three lineages of the organism (Kim *et al.*, 1999; Zhang *et al.*, 2007). While all lineages can be isolated from cattle, lineage I (LI) and lineage I/II (LI/II) strains are more frequently associated with human clinical disease than lineage II (LII) strains. Further, EHEC O157:H7 strains associated with outbreaks of severe illness in the USA have been shown to belong to a single subtype of LI/II, designated SNP clade 8 (Manning *et al.*, 2008; Laing *et al.*, 2009). Several studies have attempted to understand the genetic basis for the apparent differences in the epidemiology and virulence among EHEC O157:H7 genotypes, and recent studies have shown that LI and LI/II strains produce higher levels of Shiga toxins than bovine-associated LII strains (Zhang *et al.*, 2010). Similarly, hyper-virulent SNP clade 8 strains have been shown to have higher levels of LEE expression than EHEC O157:H7 strain Sakai from SNP clade 2 (Abu-Ali *et al.*, 2010).

A large number of factors have been shown to influence expression of LEE in both enteropathogenic *E. coli* (EPEC) and EHEC (Tree *et al.*, 2009; Kendall *et al.*, 2010). In both pathotypes, the LEE-encoded regulator (Ler) is expressed from the first LEE operon (LEE1) and is required to activate expression of all five main LEE operons (LEE1–5) (Elliott *et al.*, 2000; Sharma and Zuerner, 2004). However, at higher concentrations, Ler is considered to be repressive at LEE1 (Berdichevsky *et al.*, 2005; Yerushalmi *et al.*, 2008). To date, most of the genetic and environmental control of T3S has been shown to occur through LEE1 and Ler activation, including an activator and repressor combination within the LEE, GrlA and GrlR, which are involved in the reciprocal regulation of LEE and flagella expression (Barba *et al.*, 2005; Iyoda *et al.*, 2006; Huang and Syu, 2008; Jimenez *et al.*, 2010; Islam *et al.*, 2011;). In EPEC, LEE expression is thought to be positively regulated by the plasmid-encoded regulator C (PerC) locus and a series of PerC homologue (Pch) proteins; encoded by *pchA–E* and

pchX in EHEC O157:H7. These genes are usually associated with cryptic prophage regions encoding effector proteins that are exported by the T3S system and the Pch regulators co-ordinate the expression of these horizontally acquired effectors with the LEE-encoded T3S system (Iyoda and Watanabe, 2004; Porter *et al.*, 2005; Abe *et al.*, 2008; Yang *et al.*, 2009).

The aim of the present study was to identify OIs of EHEC O157:H7 that cross-talk with the T3S system. An OI-51 deletion mutant was shown to have lower levels of LEE expression *in vitro* and subsequent testing demonstrated that this genomic island contributes to ruminant colonization and persistence. A novel regulator, termed RgdR, was identified on OI-51 and shown to control both LEE expression and motility. The mechanism of RgdR activation of LEE was investigated.

Results

T3S screening of EHEC O-island mutants

Initial screening identified a subset of OIs with the capacity to either repress or activate T3S in EHEC strains EDL933 and TUV93-0 (Shiga toxin-negative derivative strain of EDL933). For example, TUV93-0 derived mutant's Δ OI-47, Δ OI-76 (Fig. 1A) and Δ OI-141 (Fig. 1B) all had levels of T3S above that of the wild-type parent, suggesting repression by these islands, while Δ OI-51 (Fig. 1A) and Δ OI-133 (Fig. 1B) mutants had reduced levels of T3S, suggesting activation by these islands. In the present study, we focused on the potential significance of OI-51 for colonization and how it controls T3S as variation in this region has been reported to impact on LEE regulation (Yang *et al.*, 2009).

Deletion of OI-51 leads to reduced colonization and shedding levels of *E. coli* O157 strain TUV93-0 *in vivo*

In order to assess the importance of OI-51 for colonization of ruminants, the OI-51 deletion and the parent strain were co-administered in an established ovine colonization model (Wales *et al.*, 2001). The OI-51 deletion was marked with a chloramphenicol resistance cassette and the parent strain for resistance to nalidixic acid allowing direct plating and enumeration of the two strains from faeces. Six animals were orally dosed with the two strains, with only one animal failing to be properly colonized (< 20 cfu g⁻¹ faeces from the daily counts) (Fig. 2A). For the remaining five animals the estimated relative total levels of the mutant excreted from day 5 onwards, to remove inoculum effects, were estimated with a statistically significant reduction demonstrated for the mutant ($P = 0.006$, Fig. 2B) indicating that OI-51 is important for colonization and persistence in the ruminant GI tract.

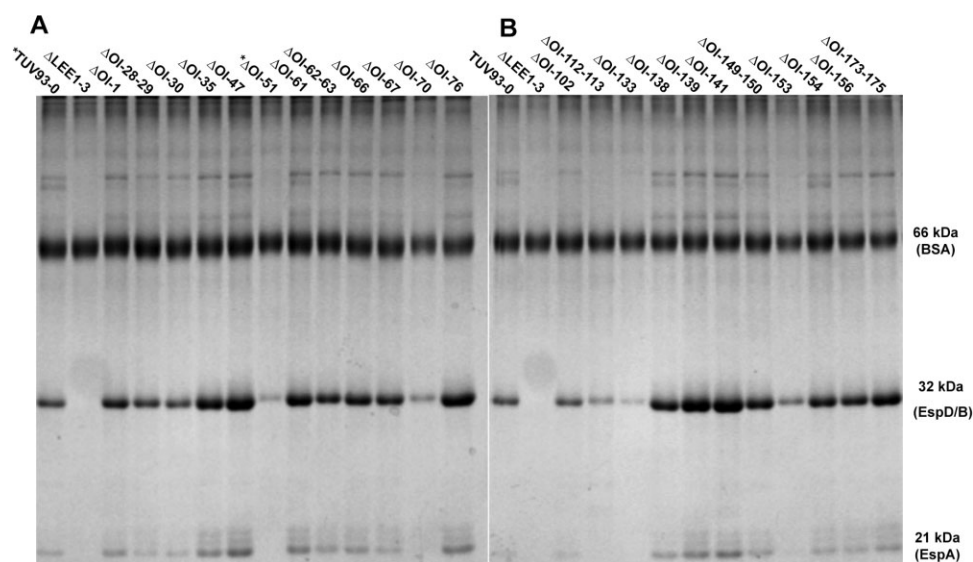


Fig. 1. Screening of EHEC O157:H7 O-island (OI) mutants for altered levels of T3S. SDS-PAGE gels showing culture supernatant secretion profiles for EHEC strain TUV93-0 and a selection of 24 isogenic deletion strains. Parent strain TUV93-0 also acted as a positive control for T3S and a T3S system mutant (Δ LEE1-3) provided a negative control for secretion. OI deletions are as defined with reference to the original designations in Perna *et al.* (2001). The translocon protein bands are indicated, EspB/D and EspA as well as BSA which was added as a loading control to rule out the precipitation process as a source of variation and to act as a co-precipitant. Strains were cultured in MEM-HEPES to an OD₆₀₀ of 1.0 and culture supernatants were TCA-precipitated, separated by SDS-PAGE and stained with Colloidal blue as described in *Experimental procedures*.

OI-51 sequence analysis

Initial phenotypic screening indicated that an EHEC O157:H7 Δ OI-51 mutant has reduced levels of T3S. OI-51 is a 14.93 kb cryptic prophage designated as CP-933C in EHEC strain EDL933 and Sp7 in EHEC strain Sakai (Hayashi *et al.*, 2001) (Fig. 3). The composition of this cryptic prophage varies in other referenced *E. coli* genomes, including *E. coli* CFT073 and ED1a (Fig. 3). Analysis of OI-51/Sp7 genomic structure shows it to be an unusual and highly degraded prophage comprised mainly of P4 phage remnants. The majority of the open reading frames annotated in OI-51/Sp7 are hypothetical although several share features with known proteins, including a P4 integrase (similar to CP4-like integrase and integrase used for 933L and LEE PAI); a P4-like excisionase (Xis); a replication gene similar to the P4 α gene; a putative DNA binding protein similar to P4 ORF88 (AlpA); a putative single-stranded DNA binding protein (ssDNA); a putative transcriptional activator similar to PerC (PchE); and phage structural genes.

ECs1581 is a positive regulator of T3S in EHEC O157:H7

Systematic analysis of cloned OI-51 regions demonstrated that a 5 kb region ($z1835$ – $z1843$) was able to rescue T3S in the TUV93-0 Δ OI-51 mutant (Fig. 4A) when

expressed from low copy number vector pWSK29 (pZ1835–43) (Table S1). Deletions across this fragment indicated that a region between $z1841$ and $z1842$ was required for this activation (data not shown). From the Sakai annotation (Fig. 3), this region does contain an open reading frame, *ecs1581*, the identical sequence being present but not annotated in the EDL933 sequence. Deletion of *ecs1581* by allelic replacement with a resistance cassette significantly reduced T3S levels in a TUV93-0 background. Further, the reduced levels of T3S in both the Δ *ecs1581* and Δ OI-51 mutants could be complemented by *in trans* expression of *ecs1581* from a low copy number plasmid (pECs1581) (Fig. 4B and C). In line with the animal colonization data, deletion of OI-51 reduced adherence of *E. coli* O157 to bovine epithelial cells and could be complemented by *ecs1581* expressed *in trans* (Fig. S1). We have therefore been able to identify *ecs1581* as a new regulator in EHEC that is able to stimulate T3S and adherence.

ECs1581 is predicted to be a small protein of 99 amino acids with a molecular mass of 11.704 kDa and a pI of 10.09. ECs1581 has no significant sequence similarity with any known protein although it does have a very short section of homology with the spacer region of Ler that bridges the known protein–protein interaction domain of Ler with its DNA binding domain (Mellies *et al.*, 2008). However, there is no significant homology of ECs1581 with these functional regions of Ler (Fig. S2).

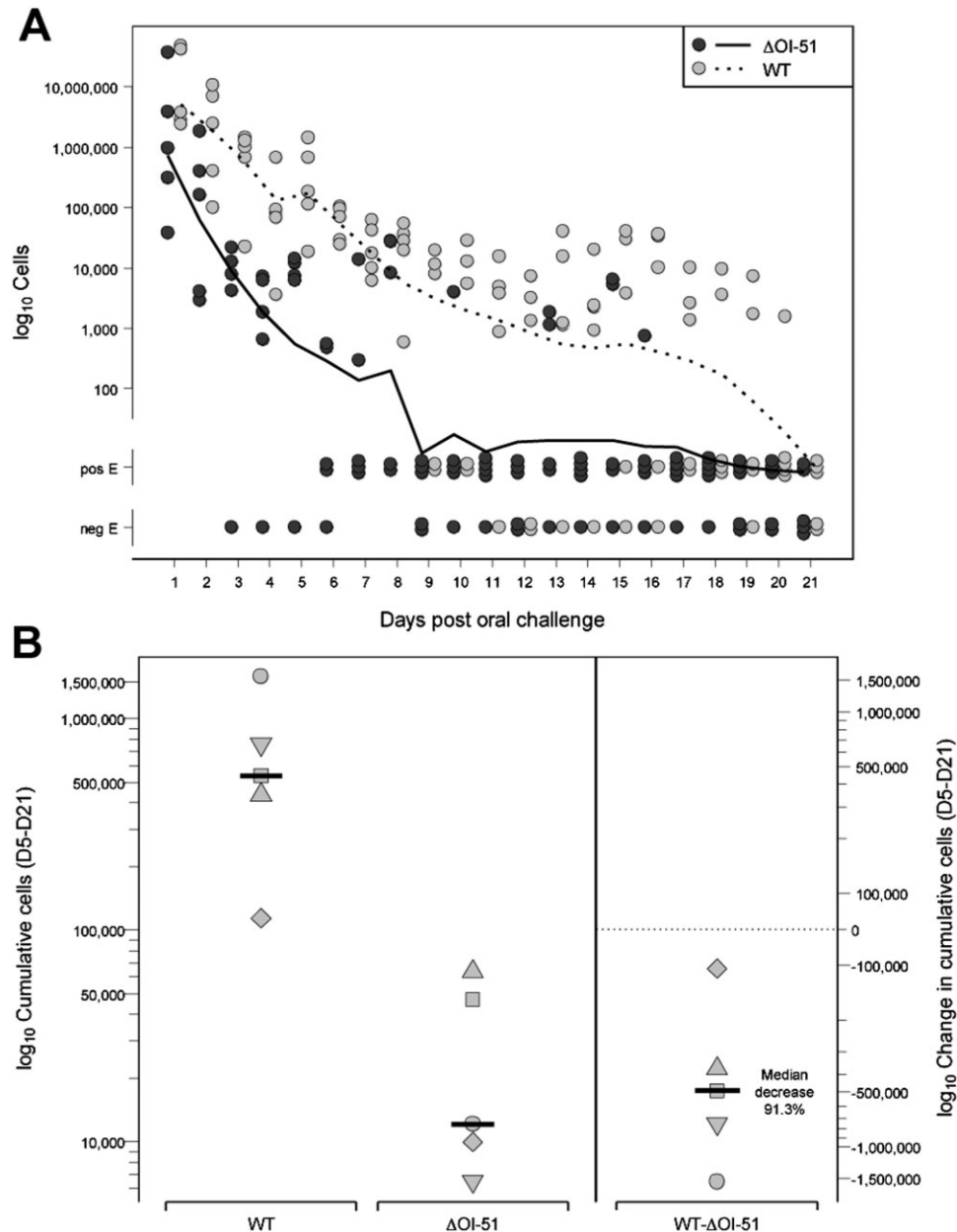


Fig. 2. OI-51 contributes to ruminant colonization. Six animals were orally dosed with both wild-type (WT) (TUV93-0) and Δ OI-51 EHEC O157:H7 strains as described in *Experimental procedures*.

A. Levels of both strains were determined daily from faecal samples. 'PosE' are samples that were positive for the strain following broth enrichment; 'NegE' samples were negative following broth enrichment.

B. The cumulative shedding levels for each animal and strain were estimated for colonized animals from day 5 onwards to avoid inoculum effects. One animal was excluded from the analysis that was not properly colonized by either strain (< 20 cfu per gram of faeces). The different symbols represent individual animals and allow direct comparison of the two strains in each animal. The right-hand panel of the graph shows the percentage mean decrease of the mutant strain by comparison with the WT. The difference in cumulative shedding levels in the five colonized animals was assessed by a one-sample *t*-test compared to a null hypothesis of 0 change. The data were log₁₀ transformed data to normalize the values, and the difference was statistically significant ($P = 0.006$).

Distribution of ECs1581 in *E. coli* strains

Sequence analysis of ECs1581 variants in different lineages of EHEC O157:H7 showed the core region of this protein (amino acids 20–68) to be highly conserved; with

distinct lineage-specific amino acid substitutions at the N- and C-termini (Fig. 5A). Interestingly, this sequence divergence was specific to LI EHEC O157 by comparison with LII and LI/II (Fig. 5A). Similar analyses of sequences immediately upstream of *ecs1581* showed divergence

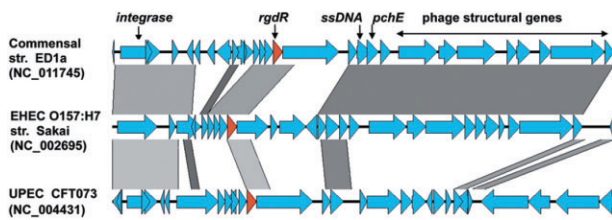


Fig. 3. Organization of OI-51 from *E. coli* O157 and nucleotide sequence homology with related regions from *E. coli* ED1a and *E. coli* CFT073. The *E. coli* O157 sequence used for the representation and analysis was from the Sakai strain and the region shown is between the chromosomal co-ordinates: 1594585–1610169 which lie between *ycdF* and *phoQ* of *E. coli* K-12 MG1655. The partially homologous region from *E. coli* CFT073 (chromosomal co-ordinates: 1377764–1393349) lies in the same chromosomal location. The related prophage in *E. coli* ED1a lies between *ycdD* and *ycdG* (chromosomal co-ordinates: 1740689–1756274). The selected regions were compared by BLASTn; regions and level of homology are indicated by the grey shading with genes encoding RgdR (ECs1581 in Sakai) and homologues (ECED1_1787 in ED1a and C1493 in CFT073) shown in orange. Other characterized genes are annotated as shown with ssDNA indicating a conserved gene in the three regions that is predicted to encode a single-stranded DNA binding protein.

between these lineages (data not shown) which could potentially impact *ecs1581* expression levels. Orthologues of ECs1581 were also found in sequenced non-O157 EHEC strains (O103:H2, O26:H11 and O111:H-); in EPEC strain O55:H7; newborn meningitis *E. coli* (NMEC) strain IHE3034; enteroaggregative *E. coli* (EAEC) strain O42; atypical EPEC strain E110019; uropathogenic *E. coli* (UPEC) strains CFT073 (O6:K2:H1) and IAI39 (O7:K1); *Shigella spp* and commensal *E. coli* strain ED1a (O81); all with > 77% homology over a region of 99 amino acids (i.e. full-length ECs1581) (Table S2). Multiple divergent copies of *ecs1581* were also present in the chromosomes of strains harbouring this regulator and these are also

detailed in Table S2. ECs1581 contains a tri-peptide arginine-glycine-aspartate (RGD) motif (amino acids 20–22, Fig. 5B) which may be important for mediating protein–protein interactions based on previously established functions of this motif (D'Souza *et al.*, 1991).

ECs1581 and orthologous protein C1493 (UPEC) can activate T3S and repress motility in EHEC and UPEC

Given the diversity of *E. coli* strains that harbour sequence-related variants of ECs1581, its prevalence and divergence among different EHEC lineages, and the fact that not all strains with this regulator harbour a LEE-encoded T3S system, we wanted to investigate the functionality of a subset of these protein variants by testing their effects on T3S and the related but more ubiquitous flagella system.

The UPEC CFT073 variant C1493 has 83% sequence similarity to ECs1581 (Fig. 5B) and is also encoded on a cryptic prophage (CP073-3) that is inserted in an identical location on the chromosome as CP-933C (OI-51/Sp7) (Mobley *et al.*, 1990; Perna *et al.*, 2001; Welch *et al.*, 2002). *c1493* was cloned into the low copy number plasmid pWSK29 to create expression plasmid pC1493 and this was assessed for its effects on T3S levels following induced expression in an EHEC *ecs1581* mutant. Expression of pC1493 rescued T3S demonstrating that the variants are functionally interchangeable, despite UPEC lineages not harbouring a LEE-encoded T3S system (Fig. 5C).

To investigate motility, a *c1493* mutant was constructed in UPEC strain CFT073 and examined in soft agar following *in trans* expression of *c1493* and *ecs1581*; in the reciprocal manner motility was examined in a EHEC

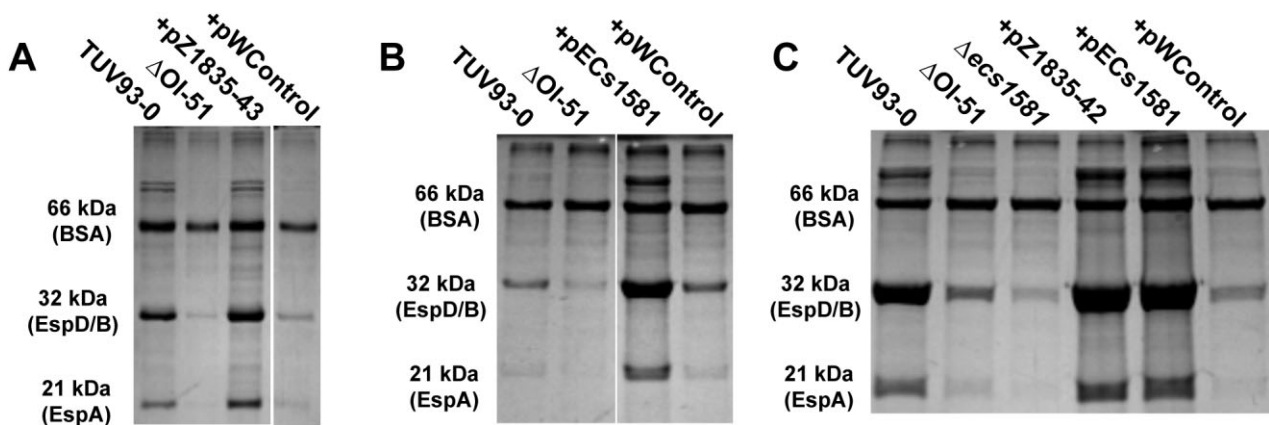


Fig. 4. Identification of *ecs1581* as a positive regulator of T3S.

A. SDS-PAGE gel showing complementation of T3S in an EHEC Δ OI-51 strain expressing a cloned OI-51 fragment containing genes *z1835*–*z1843* (pZ1835–pZ1843). The empty vector alone acted as a control (pWControl).

B. SDS-PAGE gel showing the T3S profiles for the parent strain TUV93-0 and Δ OI-51 mutant expressing pECs1581 or pWControl.

C. SDS-PAGE gel showing the T3S profiles for parent strain TUV93-0, Δ OI-51 and a Δ *ecs1581* mutant expressing pZ1835–Z1843, pECs1581 or pWControl. Protein samples were prepared and analysed as described in *Experimental procedures*.

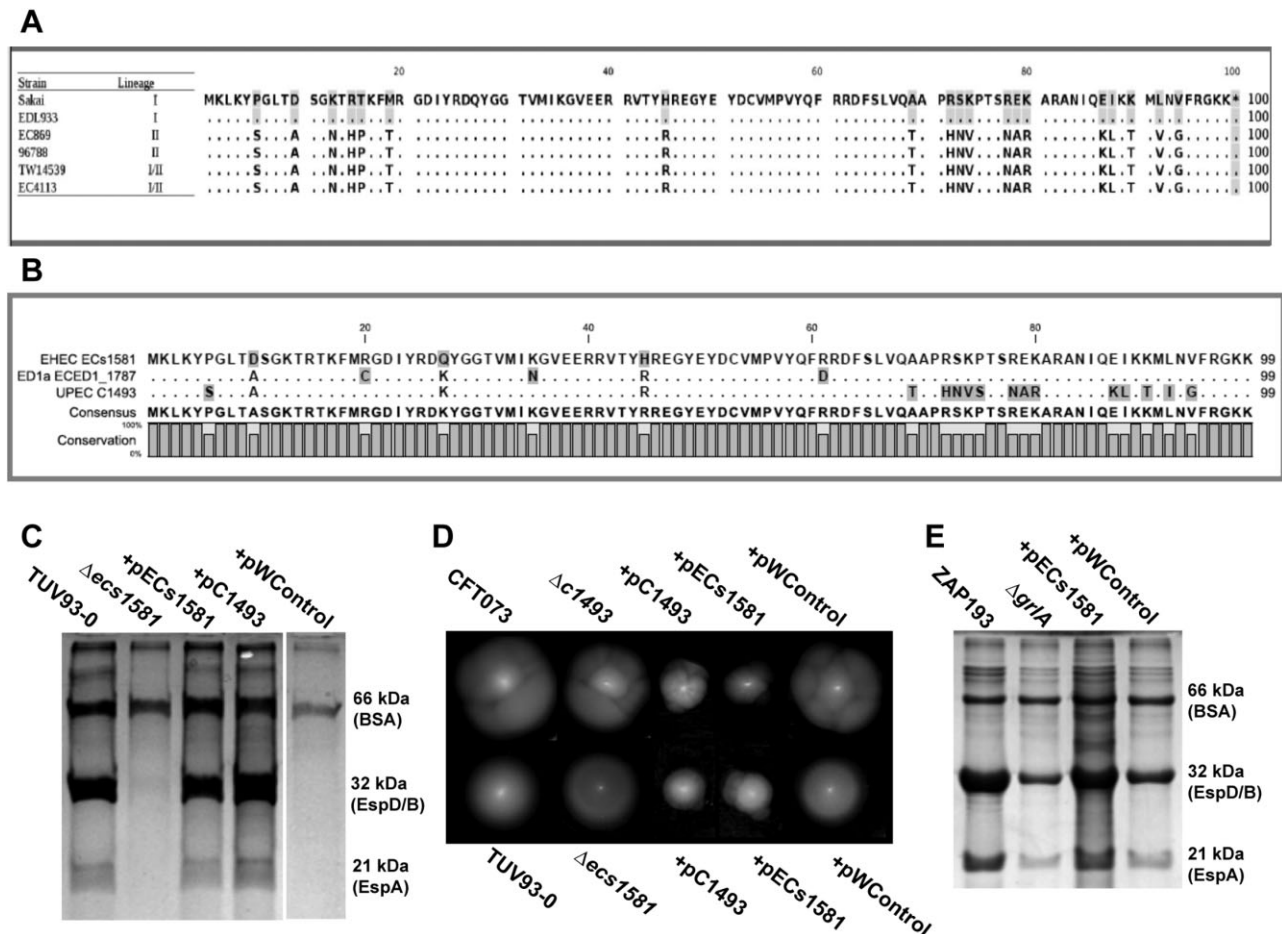


Fig. 5. Sequence comparison of ECs1581 variants in *E. coli*.

A. Multiple alignment of the predicted amino acid sequences of ECs1581 in different lineages of O157:H7. Residue divergence specific to lineage I EHEC O157:H7 isolates is marked by a pink shaded box.

B. Multiple alignment of EHEC variant ECs1581 (EDL933/Sakai) with commensal variant ECED1_1787 (ED1a) and UPEC variant C1493 (CFT073). The RGD motif is at amino acid position 20–22. Residue divergence between the strains is indicated.

C. SDS-PAGE gel showing stimulation of T3S in an EHEC $\Delta ecs1581$ mutant expressing EHEC ECs1581 (pECs1581) or UPEC C1493 (pC1493) from a low copy number plasmid. pWControl alone in the $\Delta ecs1581$ mutant was used as a negative control strain. TCA-precipitated culture supernatants were analysed as described previously in *Experimental procedures*.

D. Analysis of motility repression in an EHEC $\Delta ecs1581$ and UPEC $\Delta c1493$ deletion background expressing pECs1581, pC1493 or pWControl. Motility was assessed as described in *Experimental procedures*.

E. SDS-PAGE gel showing T3S profile for EHEC strain ZAP193 and isogenic $\Delta grlA::Tn$ mutant expressing pECs1581 or pWControl. Strains were cultured in DMEM to an OD_{600} of 1.0 and culture supernatant proteins were analysed exactly as described for strains grown in MEM-HEPES.

ecs1581 mutant strain with *in trans* expression of the two variants. Although only subtle differences in motility were observed for wild-type and mutant constructs in both EHEC and UPEC (Fig. 5D), the *in trans* and presumably higher level expression of the two variants downregulated motility, a phenotype that was not observed for the control strains expressing the empty vector (pWControl) (Fig. 5D).

In EHEC, GrlA is an established regulatory protein encoded by the LEE that positively regulates *ler* transcription and LEE gene expression; and negatively regulates motility in a feedback loop with the repressor GrlR (Iyoda *et al.*, 2006). Although unlikely given our observations in

EHEC and UPEC, we wanted to rule out the possibility that in EHEC, *ecs1581* may be acting reciprocally on T3S and flagella via *grlA*. We therefore analysed T3S and motility in a *grlA::Tn* mutant (Table S1) with and without induced expression of *ecs1581* from a low copy number plasmid. As expected, T3S levels were reduced in the *grlA::Tn* mutant compared with the parent strain (Fig. 5E). Induced expression of *ecs1581*, however, was able to increase T3S levels in the *grlA::Tn* mutant (Fig. 5E), indicating that *ecs1581* activation of T3S is mediated independently of *grlA*. Further, induced expression of *ecs1581* in the *grlA::Tn* mutant downregulated motility in this strain (data not shown). Other regulators of both T3S and motil-

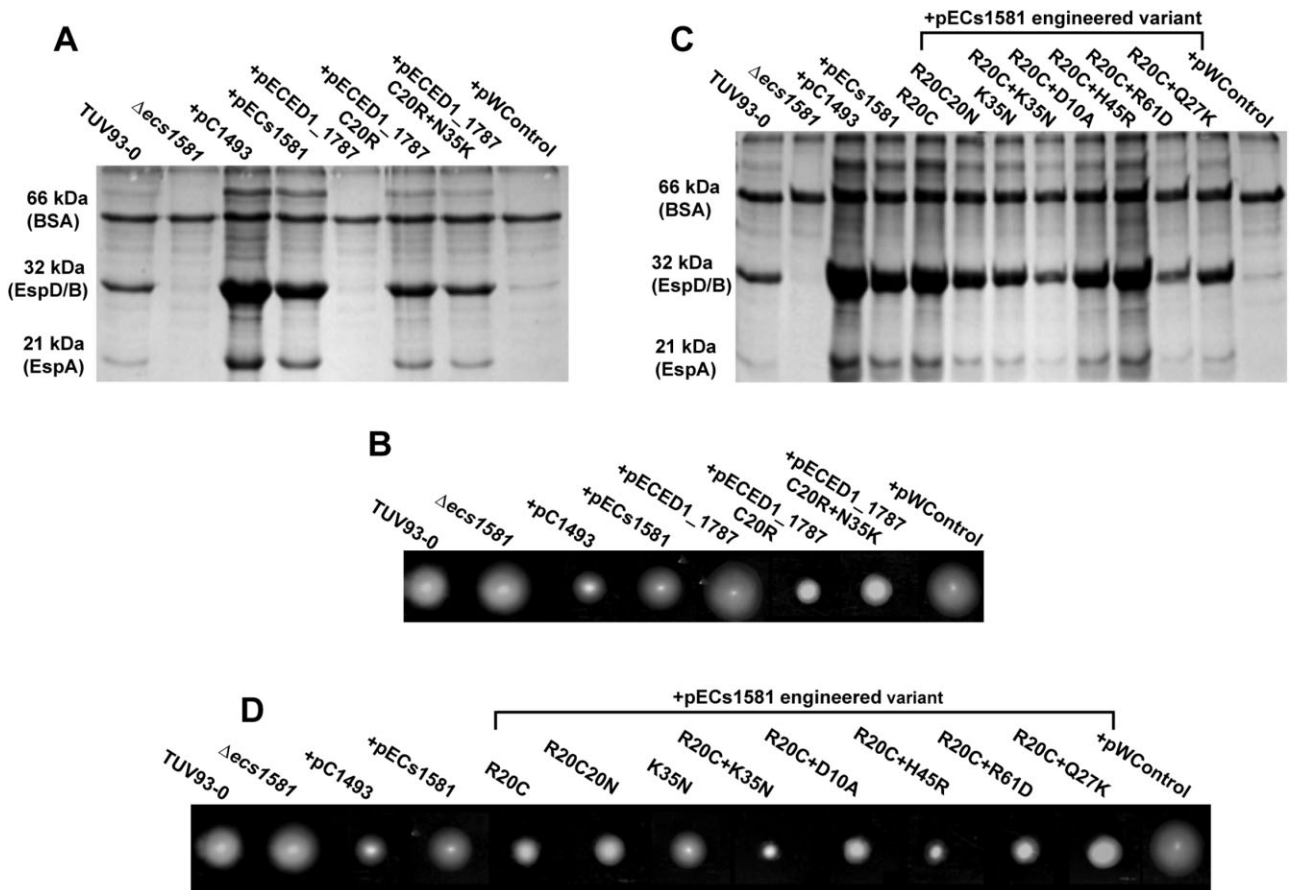


Fig. 6. Effects of natural and engineered variants of ECs1581 on T3S and motility.

A. SDS-PAGE gel showing T3S levels in EHEC strain TUV93-0 and $\Delta ecs1581$ mutant expressing wild-type and engineered variants of commensal protein ECED1_1787 from a low copy number plasmid (pECED1_1787, pECED1_1787C20R and pECED1_1787N35K). pWControl in the *ecs1581* mutant strain was used as a negative control and wild-type variants pECs1581 and pC1493 were included as a reference for high T3S levels.

B. Analysis of motility in strain TUV93-0 and $\Delta ecs1581$ mutant expressing pECED1_1787, pECED1_1787C20R, pECED1_1787N35K or pWControl.

C. SDS-PAGE gel showing T3S levels in strain TUV93-0 and $\Delta ecs1581$ mutant expressing wild-type and engineered variants of ECs1581 from a low copy number plasmid (pC1493, pECs1581, pECs1581R20C, pECs1581R20C20N, pECs1581K35N, pECs1581R20C + K35N, pECs1581R20C + D10A, pECs1581R20C + H45R, pECs1581R20C + R61D, pECs1581R20C + Q27K and pWControl).

D. Analysis of motility in the same strain set. Analysis of culture supernatants and motility was carried out as described in *Experimental procedures*.

The motility figures (B and D) are composite images from photographs of motility agar plates inoculated with the strain of interest. The same images are used for the wild types and controls in B and D.

ity include QseB-C (Hughes *et al.*, 2009; Kostakioti *et al.*, 2009), although RgdR was able to significantly increase T3S levels in a TUV93-0 $\Delta qseC$ mutant (data not shown) indicating it does not act via this pathway.

Commensal orthologue ECED1_1787 is unable to stimulate T3S or repress motility in EHEC

Commensal strain ED1a (a close relative of UPEC strain CFT073; phylogenetic group B2) also harbours an orthologue of ECs1581, ECED1_1787, which has 94% sequence similarity (Fig. 5B) (Clermont *et al.*, 2008; Jau-reguy *et al.*, 2008; Touchon *et al.*, 2009). Despite having

the highest known homology to ECs1581, even compared to other EHEC O157:H7 variants, commensal protein ECED1_1787 was unable to activate T3S or repress motility following induced expression (pECED1_1787) in an EHEC *ecs1581* mutant strain (Fig. 6A). This was informative as there are only six amino acid differences between ECs1581 and ECED1_1787 (Fig. 5B); and by the finding that C1493 has this capacity despite its lower homology to ECs1581 (83% versus 94% respectively). The residue divergence in ECED1_1787 by comparison to ECs1581 maps to the following residues: D10A, R20C, Q27K, K35N, H45R and R61D (amino acids 1–99) (Fig. 5B). Interestingly, two of these changes occur within

predicted domains with a D10A substitution in a putative N-myristylation motif (amino acids 7–12), and an R20C substitution in the EHEC-conserved RGD motif (amino acids 20–22) (Fig. 5B). Like ECED1_1787, C1493 has a D10A but harbours an RGD motif as in ECs1581 (Fig. 5B).

Amino acid substitutions in ECs1581 and ECED1_1787 alter their regulation of T3S and motility

Based on the functional differences demonstrated for the ECs1581 and ECED1_1787 proteins, a series of reciprocal substitutions in proteins ECs1581 and ECED1_1787 were constructed by site-directed mutagenesis and their effects on T3S and motility assessed in an EHEC *ecs1581* mutant. A single C20R mutation in protein ECED1_1787 (pECED1_1787C20R) was sufficient to enable this variant to now activate T3S and repress motility (Fig. 6A). An ECED1_1787 C20R + N35K double mutation (pECED1_1787C20R + N35K) did not increase T3S levels over the single C20R change (Fig. 6A). Strikingly, induced expression of ECED1_1787C20R severely repressed motility (Fig. 6B). This shift in phenotype between wild-type ECED1_1787 and engineered variant ECED1_1787C20R was intriguing as induced expression of ECED1_1787 in the same strain background, if anything, increased motility. A K35N substitution in ECs1581 (ECED1_1787N35) attenuated its stimulation of T3S in an EHEC *ecs1581* mutant when induced (pECs1581K35N) from a low copy number plasmid (Fig. 6C). Surprisingly, an R20C change in the EHEC-conserved RGD motif (ECED1_1787C20) slightly increased T3S levels following induced expression (pECs1581R20C) in this background, and this engineered variant was also more repressive for motility than wild-type ECs1581 (Fig. 6C and D). A further R20C20N mutation in this same RGD motif, however, did reduce T3S levels compared to wild-type ECs1581 (Fig. 6C). Although the small size of RgdR means that some changes could affect protein conformation and stability, these results highlight the functional importance of residue 20 and other targeted combinations for the regulatory activities of this novel regulator.

Mechanism of ECs1581 activation of T3S

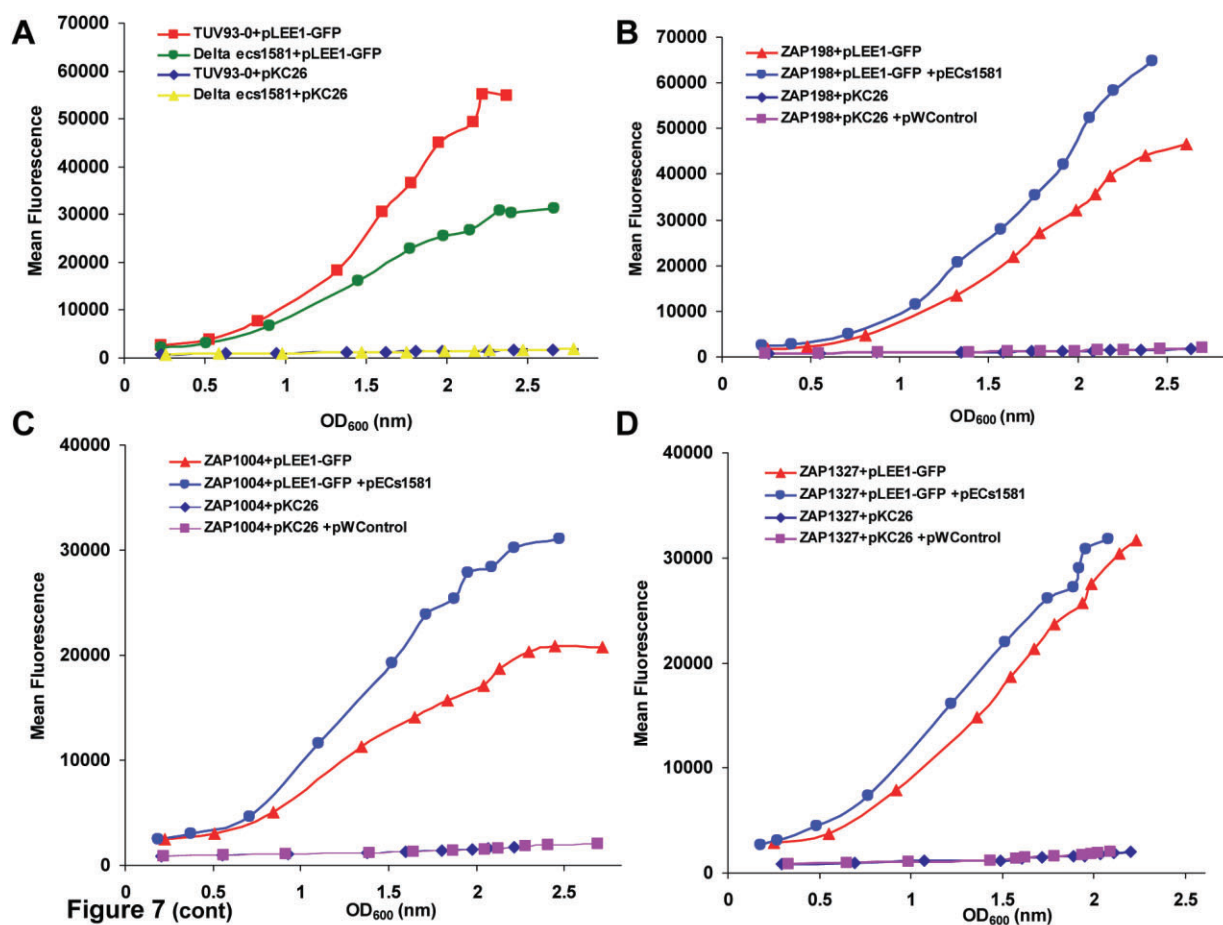
To examine how ECs1581 activates T3S, its capacity to stimulate a plasmid-encoded LEE1–GFP transcriptional fusion was studied in different genetic backgrounds. Population fluorescence levels from the LEE1 fusion were reduced significantly in an *ecs1581* mutant (Fig. 7A). *Ler*, expressed from LEE1, is required for autoregulation of LEE1 and activation of the remaining LEE operons leading to T3S. We determined whether ECs1581 was able to stimulate LEE1 expression in the absence of *ler* autoregu-

lation. Population fluorescence levels from the LEE1 reporter were determined in EHEC strains ZAP198 (WT, Table S1) and ZAP1004 (Δler , Table S1) following induction of *ecs1581* (pECs1581) (Fig. 7B and C). ECs1581 was able to stimulate LEE1 expression in the absence of *ler* (Fig. 7C). To determine if ECs1581 regulation of LEE1 expression in EHEC is direct or indirect, electrophoretic mobility shift assays (EMSA) were carried out using a 291 bp LEE1 promoter fragment (–288 to +3) and purified ECs1581. While it was evident that the protein bound to the DNA in a concentration-dependent manner (Fig. S3), ECs1581 also bound to the control DNA fragment, indicating a lack of specificity to this interaction. Of note was the finding that the commensal variant ECED1_1787, which was unable to activate T3S, was unable to bind the same LEE1 promoter fragment indicating that the regulatory activity of ECs1581 may require DNA binding (Fig. S3).

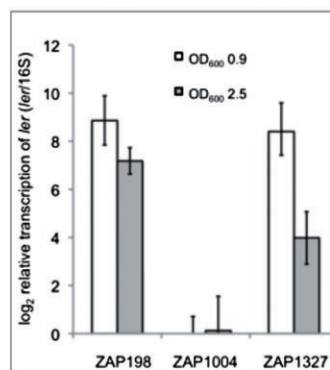
In order to confirm that ECs1581 stimulates T3S through activation of the *Ler* autoregulatory cascade, a strain (ZAP1327, Table S1) was constructed in which the LEE1 promoter was replaced by a cassette that allowed constitutive expression of *ler*, removing the normal transcriptional regulation of LEE1 including sequences required for *ler* autoregulation (Mellies *et al.*, 2008; Yerushalmi *et al.*, 2008). Constitutive *ler* transcription in this replaced promoter strain was confirmed by RT-PCR (Fig. 7E). There was, however, no detectable T3S from this strain (Fig. 7F). Induced *Ler* expression in this background was able to stimulate T3S but this was not the case for induced ECs1581. ECs1581 induction of the LEE1 reporter fusion in this strain background was detectable but remained at a constant level throughout the growth curve (Fig. 7D). This was in stark contrast to the activation kinetics measured for the LEE1 fusion in the strain with an unaltered LEE1 promoter on the chromosome. In this wild-type background, the stimulation of the LEE1 fusion increased exponentially with optical density, whereas this increase was restricted in the strain with a replaced LEE1 chromosomal promoter. Therefore, stimulation of T3S by ECs1581 occurs through initial activation of transcription from the LEE1 promoter, leading to induction of the *Ler* autoregulatory cascade that in turn promotes expression of the remaining LEE operons and T3S.

Discussion

The contribution of bacteriophages to intra- and inter-species bacterial evolution and adaptation are well documented. Enterobacteriaceae in particular are exposed to a plethora of phages in the mammalian GI tract and many of these phages carry additional genetic information that allow the host bacteria to adapt to local environments and occasionally cause disease. EHEC is an important



E



F



Fig. 7. ECs1581 stimulates T3S via LEE1.

A. Measurement of LEE1 promoter activity in EHEC strain TUV93-0 and isogenic Δ *ecs1581* mutant. A 428 bp LEE1 promoter fragment (–444 to –16 upstream from the *ler* ATG start codon) was cloned into the promoter-less green fluorescence protein (GFP) plasmid pKC26 to create transcriptional fusion plasmid pLEE1–GFP. Strains transformed with this reporter were cultured in DMEM and the fluorescence produced by each bacterial population measured every 60 min as described in the *Experimental procedures*. Corresponding OD₆₀₀ measurements were taken at each time point and plotted against the mean fluorescence intensity values for each strain. The promoter-less plasmid pKC26 in each strain background acted as a control.

B. Measurement of the pLEE1–GFP reporter and control plasmids in EHEC strain ZAP198 (Table S1) with and without induced expression of *ecs1581* (pECs1581) from low copy number plasmid pWSK29.

C. Measurement of the pLEE1–GFP reporter and control plasmids in strain ZAP198 Δ *ler* (ZAP1004, Table S1) with and without induced expression of pECs1581.

D. Measurement of the pLEE1–GFP reporter and control plasmids in strain ZAP198 with a constitutive LEE1 promoter (ZAP1327, Table S1) with and without induced expression of pECs1581.

E. RT-PCR measurement of *ler* expression levels in EHEC strains ZAP198, ZAP1004 and ZAP1327.

F. Western blot of EspD secretion by EHEC Δ *ecs1581* and the replaced LEE1 promoter mutant (ZAP1327) *trans* expressing *Ler* (pLer) or ECs1581 (pECs1581) from low copy number plasmids.

attaching and effacing pathogen that can cause serious complications in at-risk individuals due to the production of phage-encoded Shiga toxins. T3S is an essential virulence factor for EHEC which is required for successful colonization of the ruminant host. This virulence secretion system is encoded by the LEE pathogenicity island and prophage-encoded effector proteins are able to utilize this system to be exported directly into host cells to sabotage host communication pathways and promote bacterial colonization.

In this study, a number of prophage deletions (OIs) were examined for their effect on T3S levels in order to identify horizontally acquired regions that cross-talk with this essential colonization system. Of the 24 OI mutants studied in TUV93-0 (excluding the LEE1-3 deletion), five had lower levels of T3S (Fig. 1A and B), including OI-51. OI-51 is a 14.93 kb region in EDL933/TUV93-0 and 15.46 kb in the sequenced Sakai strain. Variations of this island are present in all EHEC O157:H7 strains as well as non-O157 EHEC strains, but also in specific UPEC, NMEC and commensal *E. coli* strains that do not contain a LEE-encoded T3S system. OI-51 in EDL933 contains a predicted 22 open reading frames, with a subset indicating P4 ancestry. Deletion of OI-51 reduced persistence in competitive ruminant colonization experiments indicating the potential importance of this region.

Through a combination of deletion and complementation analyses, *ecs1581* was identified as a gene able to restore T3S levels in both the OI-51 and *ecs1581* mutants. Three lineages of EHEC O157:H7 have been proposed based on genome sequences (Kim *et al.*, 1999; Zhang *et al.*, 2007). In the present study, two main sequence types of ECs1581 were identified, LI and LII & LI/II, the latter having identical protein-coding sequences. EHEC O157:H7 lineage-specific variants of ECs1581 (LI) from strains TW14539 (LI/II) and 96788 (LII) were cloned and assessed for their effects on T3S and motility. All the variants were able to induce T3S and repress motility, although slight variation in induction levels was observed (Fig. S4A and B). Sequences immediately upstream of

these two variants also differ, but it is not known if expression levels of these two variants differ as these were not assessed in this study. Variation in activity and/or expression of ECs1581 could therefore contribute to differences in T3S expression demonstrated between EHEC O157:H7 strains (Son *et al.*, 2002; Roe *et al.*, 2003; Rashid *et al.*, 2006; Yang *et al.*, 2009). We also note that genetic variability among the different EHEC O157:H7 lineages are associated with regions adjacent to *pchD* and *pchE* genes on OI-43/48 (the tellurite resistance and adhesin-conferring islands) and OI-51 respectively (Zhang *et al.*, 2007 and Yang *et al.*, 2009).

Variants of the protein from different *E. coli* strains exhibited markedly different capacities to regulate T3S. A variant present in the commensal strain ED1a did not activate T3S and contains an R20C (CGD) substitution in the EHEC-conserved RGD motif. Replacement of this residue to generate an RGD motif in the commensal variant produced a protein that now had the capacity to activate T3S. Based on this we have termed the EHEC regulator RgdR. It was interesting to note that regardless of the effect on T3S, all variants in RgdR had a marked impact on motility. In general, there was an inverse correlation between T3S and motility that was also apparent for several of the engineered variants. However, there were clear exceptions to this, for example ECs1581 R20C + K35N had a weak capacity to activate T3S yet unusually, showed the strongest capacity to inhibit motility. Additional work confirmed that this regulation of T3S and motility was independent of *grlA*, previously demonstrated to repress motility on induction of T3S (Iyoda *et al.*, 2006). RgdR is therefore identified as a novel regulator able to co-ordinate T3 and motility expression. Also of note was the finding that expression of different variants could increase bacterial binding to Congo red (Fig. S4C). Although the mechanism of this Congo red binding was not investigated in this study, recent work by Lee *et al.* (2011) has demonstrated that EDL933 is able to produce curli fibres. Taken together, the pleiotropic phenotypes displayed by RgdR including on

T3S, motility and cell surface charge suggests it may play a more global role in the cell. This would explain the presence of variants of this regulatory protein in strains without a LEE-encoded T3S system. Future work will address the regulons of RgdR in EHEC O157:H7 and strains not harbouring a LEE pathogenicity island.

RgdR was able to stimulate LEE1 expression independently of Ler and required a normal LEE1 promoter region for this activation. A construct that resulted in constitutive expression of LEE1 could be induced to secrete EspD by increased expression of Ler *in trans* but not by induced expression of RgdR *in trans*. In this background, RgdR was limited in its capacity to induce a LEE1–GFP reporter fusion construct. An explanation for this is that even though RgdR can activate the LEE1 promoter in the absence of Ler, Ler expression and the Ler auto-induction cycle is required for the activation of T3S as Ler, and not RgdR, activates transcription of LEE2-5 encoding the T3S system and secreted proteins. One explanation for the data is that ECs1581 could inhibit Ler autorepression at the LEE1 promoter. It has been proposed that high Ler levels allow binding of Ler to a low affinity site which may repress transcriptional activation. The kinetics of LEE1 activation in the presence of induced ECs1581 would fit this model. Repeated attempts to purify functional Ler to test possible interactions were unsuccessful. We note that ECs1581 does share homology with Ler over a linker region (Fig. 4D) that could be involved in this interaction.

RgdR did demonstrate a low affinity and non-specific interaction with DNA (Fig. S3), a property that was absent from the ECED_1787 variant that did not stimulate T3S. We propose that RgdR may act as a co-factor with one or more regulators of T3S (Tree *et al.*, 2009; Kendall *et al.* 2010) to control transcription from the LEE1 promoter. One regulatory network investigated was the quorum-sensing two-component system, QseBC, as this controls both T3S and motility (Hughes *et al.*, 2009; Kostakioti *et al.*, 2009). However, RgdR was able to significantly increase T3S levels in a TUV93-0 $\Delta qseC$ mutant (data not shown). Another possibility is that RgdR works with the Pch family of LEE activators (Iyoda and Watanabe, 2004; Porter *et al.*, 2005; Abe *et al.* 2008; Yang *et al.*, 2009), although there is no significant homology between the two, in particular to the region truncated in the smaller Pch variants (PchC–E) (Abe *et al.* 2008). With the exception of *pchD*, all *pch* alleles are encoded on prophage elements, and it is hypothesized that multiple prophages in the chromosome of EHEC O157:H7 may be able to recombine to affect the total copy number of *pch* genes; thus potentially impacting on levels of LEE gene expression (Ohnishi *et al.*, 2001; Iyoda and Watanabe, 2004; Asadulghani *et al.*, 2009). As with the Pch activators of LEE, all members of the RgdR family of proteins are prophage-associated. For example, EHEC strain EDL933 harbours

two RgdR truncates (both 63 amino acids), *z1197* and *z1636* that are located on OI-43 and OI-48 respectively (Table S2). In EHEC strain Sakai only one truncate exists (63 amino acids), *ecs5415*, which is located on Sakai prophage-like element 1 (SpLE1). In EHEC O26:H11 strain 11368, four copies of RgdR can be found (Table S2); three full-length (99 amino acids) variants and one truncate (63 amino acids) which is located on OI-48. Perhaps, as proposed for the Pch proteins, genome integration of CP-933C (OI-51) and similar prophages may additively influence levels of LEE gene expression in EHEC. It is therefore perhaps not surprising that a number of deleted prophage regions analysed in this study and previously (Tree *et al.*, 2011) have a regulatory impact on T3S. The RgdR, Pch and Psr families of regulators provide clear examples of how differences in the prophage repertoire of EHEC strains can impact on expression of critical colonization factors and are a major driving force in the evolution of this pathogen (Asadulghani *et al.*, 2009; Tree *et al.*, 2009; 2011).

In summary, this research has defined an important region (OI-51) necessary for EHEC O157:H7 colonization and persistence in ruminants, and has identified a completely new family of small bacterial regulators that control gene expression in *E. coli*.

Experimental procedures

Bacterial strains, plasmids and primers

All bacterial strains, plasmids and primers used in this study are detailed in Table S1.

Bacterial culture conditions and media

Bacterial strains were routinely cultured at 37°C with shaking in Luria–Bertani (LB) broth or agar. In the analysis of T3S proteins, either DMEM (Sigma-Aldrich) or MEM-HEPES (Sigma-Aldrich) supplemented with 250 nM Fe(NO₃)₂ and 0.2% glucose was used for culturing strains. When measuring promoter activity with GFP transcriptional fusions, strains were cultured in DMEM. In Congo red binding assays, CFA agar (LabM) and CFA broth [0.15% tryptone yeast extract (Becton Dickinson); 1% casamino acids (Oxoid); 0.01% Congo red (Sigma-Aldrich); 400 µM MgSO₄·7H₂O and 40 µM MnCl₂·4H₂O] was used for selecting and culturing colonies respectively. When required, Isopropyl β-D-1-thiogalactopyranoside (IPTG), 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and antibiotics were added to media at the following final concentrations: 1 mM IPTG, 20 ng ml⁻¹ X-gal, 50 µg ml⁻¹ ampicillin, 50 µg ml⁻¹ chloramphenicol, 100 µg ml⁻¹ kanamycin and 15 µg ml⁻¹ nalidixic acid.

OI-51 ruminant colonization study

Crossbred lambs (six 6-week-old) were housed in bio-secure containment level 2 accommodation and supplied with food

and water *ad libitum*. All lambs were confirmed to be free of EHEC O157 by enrichment and O157 immuno-magnetic separation prior to commencing the studies. Lambs were dosed orally with 1×10^{10} cfu of wild-type *E. coli* O157:H7 and mutant bacteria (5×10^9 cfu of each). Inocula, 10 ml resuspended in 10 ml of PBS pH 7.4, were delivered using a worming gun (Novartis Animal Health) ensuring that the whole inoculum was delivered directly to the pharynx. Rectal faecal samples from each lamb were collected daily for direct plating onto sorbitol-MacConkey (Oxoid) plates supplemented with appropriate antibiotics. Enrichment was carried out on samples in buffered peptone water for 6 h at 37°C and then plated onto sorbitol-MacConkey plates supplemented with the appropriate antibiotic. Representative colonies were confirmed to be *E. coli* O157 by latex agglutination (Oxoid). Animal experiments were performed in line with the Animals Scientific Procedures Act (1986) and were approved by the local ethical review committee.

Construction of mutations in EHEC and UPEC

Defined *ecs1581*, *c1493* and *qseC* mutants were constructed using allelic exchange as described previously by Blomfield *et al.* (1991) and Emmerson *et al.* (2006). Briefly, chromosomal regions (~600–1000 bp) flanking the genes of interest were PCR-amplified (Table S1) with a high fidelity Phusion polymerase (Fiinzyne) and cloned into a chloramphenicol-resistant temperature-sensitive exchange plasmid (pLB073) (Table S1). A kanamycin resistance cassette (Gally *et al.*, 1994) was then cloned in between the homologous flanking regions on the exchange plasmid and subjected to several rounds of temperature and antibiotic selection to achieve allelic replacement in the relevant strain background. All mutants were verified by PCR and successful complementation was achieved by providing the related genes *in trans* on low copy number plasmid pWSK29 (Table S1).

Construction of plasmids for expression and complementation

Complementation plasmids were constructed using PCR products amplified with high fidelity Phusion polymerase cloned into pWSK29 (Table S1). *E. coli* strain DH5 α was used as the intermediate host strain for cloning and all constructs were verified by sequencing.

Construction of engineered ECs1581 and ECED1_1787 protein variants

Site-directed mutagenesis was carried out using an overlap PCR approach with the primers listed in Table S1. Briefly, the 5' and 3' regions of each gene were PCR-amplified with engineered mutations in the relevant primers. The two PCR products were then used as a template for the overlap PCR using the outside primers. The PCR products were cloned into vector pWSK29, verified by sequencing (Table S1), and their effects on motility and T3S in a TUV93-0 Δ *ecs1581* mutant background assessed.

Measurement of LEE1 promoter activity

In order to assess LEE1 gene expression in both wild-type and mutant *E. coli* O157:H7 backgrounds, a 428 bp PCR-generated DNA promoter fragment (–444 to –16 bp upstream from the *ler* ATG start codon) was amplified from strain TUV93-0 using high fidelity Phusion polymerase, and cloned into the promoter-less green fluorescence protein (GFP) plasmid pKC26 to create transcriptional fusion plasmid pLEE1–GFP (Table S1). Test strains harbouring this reporter were cultured in DMEM supplemented with chloramphenicol and the GFP produced by each bacterial population was measured every 60 min by transferring 200 μ l aliquots of culture into triplicate wells in a black 96-well plate (Fluoro-Nunc) and reading the plate in a fluorimeter (FLUOstar Optima) using 485 nm absorbance and 520 nm emissions. Promoter-less plasmid pKC26 in each strain background acted as a control for auto-fluorescence. When LEE1 promoter activity was measured in response to induced expression of *ecs1581* from low copy number plasmid pWSK29 (pECs1581), strains were cultured in DMEM supplemented with chloramphenicol, ampicillin and IPTG (1 mM). All were tested in triplicate and on at least three separate occasions.

Congo red binding assay

Strains were initially streaked onto CFA agar plates (LabM) before individual colonies were taken and inoculated into CFA broth and grown overnight at 37°C with shaking. Cultures were diluted in PBS to an OD₆₀₀ of 1.0, centrifuged at 4000 g for 5 min and resuspended in 1 ml of sterile PBS. Washing was repeated a further two times to remove all traces of culture broth before finally resuspending the culture pellets in 2.5 ml of PBS containing a final concentration of 0.01% Congo red. Samples were incubated at 37°C for 30 min before a 1 ml aliquot was taken and centrifuged at 10 000 g for 10 min. Supernatants were discarded and pellets were air-dried and photographed using a Lumina digital camera. All strains were tested in triplicate and on three separate occasions.

Motility assays

Motility plates were prepared by adding 0.3% agar (Sigma) to 1% Bacto-Tryptone (BD; Becton Dickinson) broth containing 0.5% NaCl (Sigma-Aldrich). For plasmid-based complementation experiments, IPTG and ampicillin was added to plates when appropriate and all plates were poured the night before use and allowed to air-dry on the bench. Plates were stab inoculated and incubated at 37°C for 16 h. All strains were tested in triplicate and each experiment was carried out on three separate occasions.

Adherence assays

Bacteria for the assays were cultured in MEM-HEPES and added at an moi of approximately 100 bacteria to each epithelial cell. Embryonic bovine lung cells (no.21 ACC192, German Collection of Microorganisms and Cell Cultures) were cultured in MEM-HEPES in eight-well chamber slides.

The cells were washed with PBS and incubated in unsupplemented MEM-HEPES for 1 h before they were infected with bacteria. After 180 min, bacteria and EBL cells were fixed by incubating in 4% paraformaldehyde for 30 min and then washed three times with PBS. The samples were incubated for 90 min with anti-O157 LPS antibody (MAST Group) at 1:4000 and following washes incubated with Alexa Fluor 594-conjugated goat anti-rabbit immunoglobulins antibody (1/1000, Molecular Probes) for 1 h. The slides were washed and mounted (Hydromount, National Diagnostic). Imaging was carried out using a Leica fluorescence microscope and Open Lab software (PerkinElmer).

Preparation and analysis of T3S culture supernatant proteins

Bacterial strains were cultured overnight in LB broth at 37°C with shaking before being diluted to a starting optical density (OD_{600}) of 0.05 in MEM-HEPES (supplemented with glucose and iron). Cultures were grown to a final OD_{600} of 1.0 then centrifuged at 4000 *g* for 30 min at 4°C and supernatants filtered through 0.45 µm low protein-binding filters (Millipore). A 10% (v/v) final concentration of trichloroacetic acid (TCA; Sigma-Aldrich) was used to precipitate the proteins and BSA (NEB) (4 µg ml⁻¹) was added to act as a co-precipitant. Supernatants were incubated overnight at 4°C and centrifuged at 4000 *g* for 30 min at 4°C. Protein pellets were air-dried and resuspended in an appropriate volume of resuspension buffer (1.5 M Tris-HCL) in order to standardize samples and to take into account the slight variation in OD_{600} at which cultures were harvested. Culture supernatant proteins were subsequently analysed through a 12% SDS-PAGE and visualized by Colloidal blue staining (Severn Biotech). Gel images were captured using a Flowgen MultiImage light cabinet and Chemilmager 4000i v.4.04 software.

Protein detection

SDS-PAGE separated proteins were transferred onto Hybond ECL nitrocellulose membrane (Amersham Biosciences) using a Trans-Blot electrophoretic transfer cell (Bio-Rad). Nitrocellulose membranes were blocked with 8% (w/v) dried milk powder (Marvel) in PBS at 4°C overnight and incubated with the relevant antibodies diluted in wash buffer [1% dried milk (Marvel) and 0.05% (v/v) polyoxyethylenesorbitan monolaurate (Tween 20, Sigma-Aldrich) in PBS] at the following dilutions: mouse monoclonal anti-EspD (kindly provided by Professor T. Chakraborty, University of Giessen) and rabbit polyclonal anti-mouse IgG HRP conjugated antibodies (Dako) were both diluted 1:4000. All antibody-membrane incubations were carried out for 1 h at RT on a platform shaker and were washed for 3 × 10 min in wash buffer before and after each antibody step. For ECL detection, membranes were incubated in 2.5 ml of ECL Solution 1 (Amersham) mixed with 2.5 ml of ECL Solution 2 (Amersham) for 5 min at RT on the bench. Chemiluminescence was detected on Hyperfilm ECL chemiluminescence film (Amersham) developed in a Protec automatic film processor (Optimax). Images were taken using a Flowgen MultiImage light cabinet and Chemilmager 4000i v.4.04 software.

Native protein purification

Escherichia coli BL21 (DE3) (Invitrogen) cells harbouring plasmid pET28a-ECs1581, pET28a-ECED1_1787 or pET28a-ECED1_1787C20R were grown in LB broth in the presence of kanamycin at 37°C with shaking to an OD_{600} of 0.5. Cultures were induced with IPTG and grown for a further 4 h under the same conditions before harvesting by centrifugation at 4000 *g* for 30 min. Cell pellets were suspended in lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole and 1 mg ml⁻¹ lysozyme, pH 8.0) and lysed by sonication (6 × 10 s pulses with a 10 s cooling period on ice in between each pulse). Lysates were centrifuged at 10 000 *g* at 4°C for 30 min, incubated with Ni²⁺ NT resin (Qiagen) for 1 h at RT on a platform shaker and loaded onto a purification column (Bio-Rad). The column was washed twice with wash buffer (50 mM NaH₂PO₄, 300 mM NaCl, 20 mM imidazole, pH 8.0) and the protein was eluted from the column in elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 250 mM imidazole, pH 8.0). Crude preps were transferred into dialysis tubing (8000 kDa molecular weight cut-off, Medicell international) and dialysed in 2 l of sterile PBS which was replenished every 24 h for a total of three changes. Total protein concentration was determined by BCA according to the manufacturer's instructions (Novagen) and purity was verified by SDS-PAGE and antibody detection of recombinant His-tagged proteins. Purified proteins were subsequently used in electrophoretic mobility shift and *in vitro* transcription assays along with PCR generated DNA promoter fragments.

Electrophoretic mobility shift assay

To determine if ECs1581 activation of LEE1 is direct or indirect, a 291 bp LEE1 promoter (−288 to +3, with respect to the *ler* ATG start codon) and a control *gapA* fragment (also 291 bp) were PCR-amplified (Table S1) from EHEC strain TUV93-0 and subjected to EMSA with either purified ECs1581, ECED1_1787 or ECED1_1787C20R proteins. EMSA was carried out as described previously (Tree *et al.*, 2011). Briefly, the purified proteins were co-incubated with concentrations of ddUTP-11-DIG (Roche) end-labelled DNA in binding buffer [10 mM Tris (pH 8.0), 50 mM KCl, 5 mM EDTA, 200 µg ml⁻¹ BSA, 100 ng ml⁻¹ poly d(I-C), 50 mM glutamate and 5% glycerol] for 30 min at 25°C and loaded onto 5% non-denaturing polyacrylamide gels in 0.5X TBE. Fifty nanograms of unlabelled DNA was used for competition experiments and added to the binding reaction. The DNA and DNA-protein complexes were electro-transferred onto nylon membranes and developed with AP-conjugated anti-DIG antibody (Roche) as per the manufacturer's instructions.

RT-qPCR

Total RNA was purified from the bacteria using an RNeasy kit (Qiagen) and random primers used to reverse transcribe the RNA using Affinityscript (Stratagene). qPCR was carried out with a PowerSybr mastermix (Applied Biosystems) and amplified in a MxPro 3000 qPCR machine (Stratagene). The qPCR primers used are listed in Table S1. Transcript abundance was normalized to 16S rRNA and relative transcription calculated using MxPro software (Stratagene).

Bioinformatic analyses

Variants of ECs1581 were identified using BLASTp (<http://www.ncbi.nlm.nih.gov>) and sequences were obtained from GeneBank (<http://www.ncbi.nlm.nih.gov/GenBank>). Putative hypothetical domains in ECs1581 were identified using MyHits (<http://myhits.isb-sib.ch/cgi-bin/index>) and KEGG (<http://www.genome.jp/kegg>). OI-51 prophage alignments were generated using Easyfig (<http://easyfig.sourceforge.net/>). Sequence alignments were performed using ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2>), EMBOSS Align (<http://www.ebi.ac.uk/Tools/emboss/align>) and LALIGN (http://www.ch.embnet.org/software/LALIGN_form.html). CLC Sequence Viewer (<http://www.clcbio.com>) was used to construct and visualize alignments among the protein-coding regions of ECs1581; the ECs1581 region was identified in un-annotated draft *E. coli* sequences through BLASTn searches.

Sequencing

All DNA sequencing was carried out by GenePool, School of Biological Sciences, University of Edinburgh, and was performed using an ABI Prism BigDye terminator cycle sequencing kit version 3.1 (Applied Biosystems). Reactions were analysed on an ABI 3730 DNA sequencer.

Acknowledgements

A.F.F. was supported by a studentship from the College of Medicine and Veterinary Medicine at the University of Edinburgh. X.X. was supported by a studentship from the China Scholarship Council. D.L.G., J.J.T. and S.P.M. would like to acknowledge funding from DEFRA under the Veterinary Training and Research Initiative (VT0102), core support from the BBSRC at the Roslin Institute and Wellcome Trust grant 090231/Z/09/Z. Salary (R.R. and T.S.) and consumable support for this work were provided by the Department for Employment and Learning (Northern Ireland) through its 'Strengthening the all-island Research Base' initiative. V.G. and C.L. were funded by the Public Health Agency of Canada, Alberta Innovates Technology Futures and the Natural Sciences and Engineering Research Council of Canada. M.K. was supported by the Darwin Trust of Edinburgh. We would like to thank Yongxiang Zhang for helpful discussions; Professor Erik Denamur for kindly providing commensal *E. coli* strain ED1a; Professor Trinad Chakraborty for his generous supply of monoclonal antibodies; and Dr Kenneth Campellone at UC Berkeley for his help constructing the OI deletion collection.

References

- Abe, H., Miyahara, A., Oshima, T., Tashiro, K., Ogura, Y., Kuhara, S., *et al.* (2008) Global regulation by horizontally transferred regulators establishes the pathogenicity of *Escherichia coli*. *DNA Res* **15**: 25–38.
- Abu-Ali, G.S., Ouellette, L.M., Henderson, S.T., Whittam, T.S., and Manning, S.D. (2010) Differences in adherence and virulence gene expression between two outbreak strains of enterohaemorrhagic *Escherichia coli* O157:H7. *Microbiology* **156**: 408–419.
- Armstrong, G.L., Hollingsworth, J., and Morris, J.G., Jr (1996) Emerging foodborne pathogens: *Escherichia coli* O157:H7 as a model of entry of a new pathogen into the food supply of the developed world. *Epidemiol Rev* **18**: 29–51.
- Asadulghani, M., Ogura, Y., Ooka, T., Itoh, T., Sawaguchi, A., Iguchi, A., *et al.* (2009) The defective prophage pool of *Escherichia coli* O157: prophage-prophage interactions potentiate horizontal transfer of virulence determinants. *PLoS Pathog* **5**: e1000408.
- Barba, J., Bustamante, V.H., Flores-Valdez, M.A., Deng, W., Finlay, B.B., and Puente, J.L. (2005) A positive regulatory loop controls expression of the locus of enterocyte effacement-encoded regulators Ler and GrlA. *J Bacteriol* **187**: 7918–7930.
- Berdichevsky, T., Friedberg, D., Nadler, C., Rokney, A., Oppenheim, A., and Rosenshine, I. (2005) Ler is a negative autoregulator of the LEE1 operon in enteropathogenic *Escherichia coli*. *J Bacteriol* **187**: 349–357.
- Besser, R.E., Griffin, P.M., and Slutsker, L. (1999) *Escherichia coli* O157:H7 gastroenteritis and the hemolytic uremic syndrome: an emerging infectious disease. *Annu Rev Med* **50**: 355–367.
- Blomfield, I.C., Vaughn, V., Rest, R.F., and Eisenstein, B.I. (1991) Allelic exchange in *Escherichia coli* using the *Bacillus subtilis* sacB gene and a temperature-sensitive pSC101 replicon. *Mol Microbiol* **5**: 1447–1457.
- Campellone, K.G., Robbins, D., and Leong, J.M. (2004) EspFU is a translocated EHEC effector that interacts with Tir and N-WASP and promotes Nck-independent actin assembly. *Dev Cell* **7**: 217–228.
- Caprioli, A., Morabito, S., Brugere, H., and Oswald, E. (2005) Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. *Vet Res* **36**: 289–311.
- Chase-Topping, M., Gally, D., Low, C., Matthews, L., and Woolhouse, M. (2008) Super-shedding and the link between human infection and livestock carriage of *Escherichia coli* O157. *Nat Rev Microbiol* **6**: 904–912.
- Chaudhuri, R.R., Sebaihia, M., Hobman, J.L., Webber, M.A., Leyton, D.L., Goldberg, M.D., *et al.* (2010) Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain 042. *PLoS ONE* **5**: e8801.
- Clermont, O., Lescat, M., O'Brien, C.L., Gordon, D.M., Tenaillon, O., and Denamur, E. (2008) Evidence for a human-specific *Escherichia coli* clone. *Environ Microbiol* **10**: 1000–1006.
- D'Souza, S.E., Ginsberg, M.H., and Plow, E.F. (1991) Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *Trends Biochem Sci* **16**: 246–250.
- Elliott, S.J., Sperandio, V., Giron, J.A., Shin, S., Mellies, J.L., Wainwright, L., *et al.* (2000) The locus of enterocyte effacement (LEE)-encoded regulator controls expression of both LEE- and non-LEE-encoded virulence factors in enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect Immun* **68**: 6115–6126.
- Emmerson, J.R., Gally, D.L., and Roe, A.J. (2006) Generation of gene deletions and gene replacements in *Escherichia coli* O157:H7 using a temperature sensitive allelic exchange system. *Biol Proced Online* **8**: 153–162.
- Gally, D.L., Rucker, T., and Blomfield, I.C. (1994) The

- Leucine-responsive regulatory protein binds to the *fim* switch to control phase variation of type 1 fimbriae expression in *Escherichia coli* K-12. *J Bacteriol* **176**: 5665–5672.
- Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., *et al.* (2001) Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* **8**: 11–22.
- Hemrajani, C., Berger, C.N., Robinson, K.S., Marches, O., Mousnier, A., and Frankel, G. (2010) NleH effectors interact with Bax inhibitor-1 to block apoptosis during enteropathogenic *Escherichia coli* infection. *Proc Natl Acad Sci USA* **107**: 3129–3134.
- Holden, N., Totsika, M., Dixon, L., Catherwood, K., and Gally, D.L. (2007) Regulation of P-fimbrial phase variation frequencies in *Escherichia coli* CFT073. *Infect Immun* **75**: 3325–3334.
- Huang, L.H., and Syu, W.J. (2008) GrlA of enterohemorrhagic *Escherichia coli* O157:H7 activates LEE1 by binding to the promoter region. *J Microbiol Immunol Infect* **41**: 9–16.
- Hughes, D.T., Clarke, M.B., Yamamoto, K., Rasko, D.A., and Sperandio, V. (2009) The QseC adrenergic signaling cascade in Enterohemorrhagic *E. coli* (EHEC). *PLoS Pathog* **5**: e1000553.
- Islam, M.S., Bingle, L.E., Pallen, M.J., and Busby, S.J. (2011) Organization of the LEE1 operon regulatory region of enterohaemorrhagic *Escherichia coli* O157:H7 and activation by GrlA. *Mol Microbiol* **79**: 468–483.
- Iyoda, S., and Watanabe, H. (2004) Positive effects of multiple *pch* genes on expression of the locus of enterocyte effacement genes and adherence of enterohaemorrhagic *Escherichia coli* O157:H7 to HEp-2 cells. *Microbiology* **150**: 2357–2571.
- Iyoda, S., Koizumi, N., Satou, H., Lu, Y., Saitoh, T., Ohnishi, M., and Watanabe, H. (2006) The GrlR-GrlA regulatory system co-ordinately controls the expression of flagellar and LEE-encoded type III protein secretion systems in enterohemorrhagic *Escherichia coli*. *J Bacteriol* **188**: 5682–5692.
- Jauregui, F., Landraud, L., Passet, V., Diancourt, L., Frapy, E., Guigon, G., *et al.* (2008) Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics* **9**: 560.
- Jimenez, R., Cruz-Migoni, S.B., Huerta-Saquero, A., Bustamante, V.H., and Puente, J.L. (2010) Molecular characterization of GrlA, a specific positive regulator of *ler* expression in enteropathogenic *Escherichia coli*. *J Bacteriol* **18**: 4627–4642.
- Kaper, J.B., Nataro, J.P., and Mobley, H.L. (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol* **2**: 123–140.
- Karmali, M.A. (2004) Infection by Shiga toxin-producing *Escherichia coli*: an overview. *Mol Biotechnol* **26**: 117–122.
- Kendall, M.M., Rasko, D.A., and Sperandio, V. (2010) The LysR-type regulator QseA regulates both characterized and putative virulence genes in enterohaemorrhagic *Escherichia coli* O157: H7. *Mol Microbiol* **76**: 1306–1321.
- Kenny, B., Abe, A., Stein, M., and Finlay, B.B. (1997) Enteropathogenic *Escherichia coli* protein secretion is induced in response to conditions similar to those in the gastrointestinal tract. *Infect Immun* **65**: 2606–2612.
- Kim, J., Nietfeldt, J., and Benson, A.K. (1999) Octamer-based genome scanning distinguishes a unique subpopulation of *Escherichia coli* O157:H7 strains in cattle. *Proc Natl Acad Sci USA* **96**: 13288–13293.
- Kostakioti, M., Hadjifrangiskou, M., Pinkner, J.S., and Hultgren, S.J. (2009) QseC-mediated dephosphorylation of QseB is required for expression of genes associated with virulence in uropathogenic *Escherichia coli*. *Mol Microbiol* **73**: 1020–1031.
- La Ragione, R.M., Best, A., Woodward, M.J., and Wales, A.D. (2009) *Escherichia coli* O157:H7 colonization in small domestic ruminants. *FEMS Microbiol Rev* **33**: 394–410.
- Laing, C.R., Buchanan, C., Taboada, E.N., Zhang, Y., Karmali, M.A., Thomas, J.E., and Gannon, V.P. (2009) In silico genomic analyses reveal three distinct lineages of *Escherichia coli* O157:H7, one of which is associated with hyper-virulence. *BMC Genomics* **10**: 287.
- Lee, J.H., Kim, Y.G., Cho, M.H., Wood, T.K., and Lee, J. (2011) Transcriptomic analysis for genetic mechanisms of the factors related to biofilm formation in *Escherichia coli* O157:H7. *Curr Microbiol* **62**: 1321–1330.
- Low, A.S., Dziva, F., Torres, A.G., Martinez, J.L., Rosser, T., Naylor, S., *et al.* (2006) Cloning, expression, and characterization of fimbrial operon F9 from enterohemorrhagic *Escherichia coli* O157:H7. *Infect Immun* **74**: 2233–2244.
- Manning, S.D., Motiwala, A.S., Springman, A.C., Qi, W., Lacher, D.W., Ouellette, L.M., *et al.* (2008) Variation in virulence among clades of *Escherichia coli* O157 : H7 associated with disease outbreaks. *Proc Natl Acad Sci USA* **105**: 4868–4873.
- Mellies, J.L., Larabee, F.J., Zarr, M.A., Horback, K.L., Lorenzen, E., and Mavor, D. (2008) Ler interdomain linker is essential for anti-silencing activity in enteropathogenic *Escherichia coli*. *Microbiology* **154**: 3624–3638.
- Mobley, H.L., Green, D.M., Trifillis, A.L., Johnson, D.E., Chippendale, G.R., Lockett, C.V., *et al.* (1990) Pylonephritogenic *Escherichia coli* and killing of cultured human renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect Immun* **58**: 1281–1289.
- Moriel, D.G., Bertoldi, I., Spagnuolo, A., Marchi, S., Rosini, R., Nesta, B., *et al.* (2010) Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* **107**: 9072–9077.
- Nataro, J.P., and Kaper, J.B. (1998) Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**: 142–201.
- Naylor, S.W., Roe, A.J., Nart, P., Spears, K., Smith, D.G., Low, J.C., and Gally, D.L. (2005) *Escherichia coli* O157:H7 forms attaching and effacing lesions at the terminal rectum of cattle and colonization requires the LEE4 operon. *Microbiology* **151**: 2773–2781.
- Newton, H.J., Pearson, J.S., Badea, L., Kelly, M., Lucas, M., Holloway, G., *et al.* (2010) The type III effectors NleE and NleB from enteropathogenic *E. coli* and OspZ from *Shigella* block nuclear translocation of NF-kappaB p65. *PLoS Pathog* **6**: e1000898.
- Nie, H., Yang, F., Zhang, X., Yang, J., Chen, L., Wang, J., *et al.* (2006) Complete genome sequence of *Shigella flexneri* 5b and comparison with *Shigella flexneri* 2a. *BMC Genomics* **7**: 173.
- Ogura, Y., Ooka, T., Iguchi, A., Toh, H., Asadulghani, M.,

- Oshima, K., *et al.* (2009) Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *Proc Natl Acad Sci USA* **106**: 17939–17944.
- Ohnishi, M., and Hayashi, T. (2002) [Genetic diversity of enterohemorrhagic *Escherichia coli*]. *Nihon Rinsho* **60**: 1077–1082.
- Ohnishi, M., Kurokawa, K., and Hayashi, T. (2001) Diversification of *Escherichia coli* genomes: are bacteriophages the major contributors? *Trends Microbiol* **9**: 481–485.
- Perna, N.T., Plunkett, G., Burland, V., Mau, B., Glasner, J.D., Rose, D.J., *et al.* (2001) Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**: 529–533.
- Porter, M.E., Mitchell, P., Free, A., Smith, D.G., and Gally, D.L. (2005) The LEE1 promoters from both enteropathogenic and enterohemorrhagic *Escherichia coli* can be activated by PerC-like proteins from either organism. *J Bacteriol* **187**: 458–472.
- Rashid, R.A., Tabata, T.A., Oatley, M.J., Besser, T.E., Tarr, P.I., and Moseley, S.L. (2006) Expression of putative virulence factors of *Escherichia coli* O157:H7 differs in bovine and human infections. *Infect Immun* **74**: 4142–4148.
- Rasko, D., Rosovitz, M., Myers, G., Seshadri, R., Cer, R., Jiang, L., *et al.* (2008) Direct submission. J. Craig Venter Institute, 9704 Medical Centre, Rockville, MD 20850, USA.
- Roe, A.J., Hoey, D.E., and Gally, D.L. (2003) Regulation, secretion and activity of type III-secreted proteins of enterohaemorrhagic *Escherichia coli* O157. *Biochem Soc Trans* **31**: 98–103.
- Sharma, V.K., and Zuercher, R.L. (2004) Role of hha and ler in transcriptional regulation of the esp operon of enterohemorrhagic *Escherichia coli* O157:H7. *J Bacteriol* **186**: 7290–7301.
- Son, W.G., Graham, T.A., and Gannon, V.P. (2002) Immunological characterization of *Escherichia coli* O157:H7 intimin gamma1. *Clin Diagn Lab Immunol* **9**: 46–53.
- Sperandio, V., Li, C.C., and Kaper, J.B. (2002) Quorum-sensing *Escherichia coli* regulator A: a regulator of the LysR family involved in the regulation of the locus of enterocyte effacement pathogenicity island in enterohemorrhagic *E. coli*. *Infect Immun* **70**: 3085–3093.
- Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., *et al.* (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* **5**: e1000344.
- Tree, J.J., Wolfson, E.B., Wang, D., Roe, A.J., and Gally, D.L. (2009) Controlling injection: regulation of type III secretion in enterohaemorrhagic *Escherichia coli*. *Trends Microbiol* **17**: 361–370.
- Tree, J.J., Roe, A.J., Flockhart, A., McAteer, S.P., Xu, X., Mahajan, A., *et al.* (2011) Transcriptional regulators of the GAD acid stress island are carried by effector protein-encoding prophages and indirectly control type III secretion in enterohaemorrhagic *Escherichia coli* O157:H7. *Mol Microbiol* **80**: 1349–1365.
- Tupper, A.E., Owen-Hughes, T.A., Ussery, D.W., Santos, D.S., Ferguson, D.J., Sidebotham, J.M., *et al.* (1994) The chromatin-associated protein H-NS alters DNA topology in vitro. *EMBO J* **13**: 258–268.
- Wales, A.D., Clifton-Hadley, F.A., Cookson, A.L., Dibb-Fuller, M.P., La Ragione, R.M., Spriggs, K.A., *et al.* (2001) Experimental infection of six-month-old sheep with *Escherichia coli* O157:H7. *Vet Rec* **148**: 630–631.
- Welch, R.A., Burland, V., Plunkett, G., 3rd, Redford, P., Roesch, P., Rasko, D., *et al.* (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* **99**: 17020–17024.
- Yang, Z., Kim, J., Zhang, C., Zhang, M., Nietfeldt, J., Southward, C.M., *et al.* (2009) Genomic instability in regions adjacent to a highly conserved *pch* prophage in *Escherichia coli* O157:H7 generates diversity in expression patterns of the LEE pathogenicity island. *J Bacteriol* **191**: 3553–3568.
- Ye, C., Lan, R., Xia, S., Zhang, J., Sun, Q., Zhang, S., *et al.* (2010) Emergence of a new multidrug-resistant serotype X variant in an epidemic clone of *Shigella flexneri*. *J Clin Microbiol* **48**: 419–426.
- Yerushalmi, G., Nadler, C., Berdichevski, T., and Rosenshine, I. (2008) Mutational analysis of the locus of enterocyte effacement-encoded regulator (Ler) of enteropathogenic *Escherichia coli*. *J Bacteriol* **190**: 7808–7818.
- Zhang, Y., Laing, C., Steele, M., Ziebell, K., Johnson, R., Benson, A.K., *et al.* (2007) Genome evolution in major *Escherichia coli* O157:H7 lineages. *BMC Genomics* **8**: 121.
- Zhang, Y., Laing, C., Zhang, Z., Hallewell, J., You, C., Ziebell, K., *et al.* (2010) Lineage and host-source are both correlated with levels of Shiga toxin 2 production by *Escherichia coli* O157:H7 strains. *Appl Environ Microbiol* **76**: 474–482.
- Zhou, Z., Li, X., Liu, B., Beutin, L., Xu, J., Ren, Y., *et al.* (2010) Derivation of *Escherichia coli* O157:H7 from its O55:H7 precursor. *PLoS ONE* **5**: e8700.

Supporting information

Additional supporting information may be found in the online version of this article.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.